

# TRANSLATION

## (12) INTERNATIONAL APPLICATION PUBLISHED UNDER PATENT COOPERATION TREATY (PCT)

(19) World Organization for Intellectual Property  
International Office

(43) Date of international publication  
August 30, 2001

PCT

(10) International Publication No.  
WO 01/62781 A2

- |  |   |
|--|---|
| <p>(51) International Classification<sup>7</sup>: C07K 14/415</p> <p>(21) International file No.: PCT/EP01/01720</p> <p>(22) Date of international filing: February 16, 2001</p> <p>(25) Filing language: German</p> <p>(26) Publication language: German</p> <p>(30) Priority details:<br/>100 09 002.8 February 25, 2000 DE</p> <p>(71) Applicant (<i>for all designated states except U.S.</i>):<br/>SUNGENE GMBH &amp; CO. KGaA (DE/DE);<br/>06468 Gatersleben (DE)</p> <p>(72) Inventors; and</p> <p>(75) Inventors/Applicants (<i>only for U.S.</i>): Karin<br/>HERBERS (DE/DE); Am Hange 6, 06484<br/>Quedlinburg (DE). Ralf BADUR (DE/DE); Von<br/>Eckenbrecher Weg 2, 38642 Goslar (DE). Irene<br/>KUNZE (DE/DE); Mühlenweg 11, 06466<br/>Gatersleben (DE). Susanne SOMMER<br/>(DE/DE); Brechtstr. 8, 06484 Quedlinburg (DE).<br/>Rainer LEMKE (DE/DE); Halberstaedter Str.<br/>14, 06484 Quedlinburg (DE). Michael GEIGER<br/>(DE/DE); Neuer Weg 15, 06484 Quedlinburg<br/>(DE).</p> | <p>(74) Attorney: Andreas BIEBERBACH; BASF<br/>Aktiengesellschaft, 67056 Ludwigshafen (DE).</p> <p>(81) Designated States (national): AE, AG, AL, AM,<br/>AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA,<br/>CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE,<br/>ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL,<br/>IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR,<br/>LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,<br/>MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE,<br/>SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG,<br/>US, UZ, VN, YU, ZA, ZW.</p> <p>(84) Designated States (regional): ARIPO Patent<br/>(GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ,<br/>UG, ZW), Eurasian Patent (AM, AZ, BY, KG,<br/>KZ, MD, RU, TJ, TM), European Patent (AT,<br/>BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,<br/>IT, LU, MC, NL, PT, SE, TR), OAPI Patent (BF,<br/>BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR,<br/>NE, SN, TD, TG).</p> <p><b>Published</b><br/>– <i>Without an International Search Report and to<br/>be republished after receipt of the Report</i></p> <p><i>For an explanation of the two-letter codes and the other<br/>abbreviations, please refer to the Guidance Notes on<br/>Codes and Abbreviations at the beginning of each<br/>regular issue of the PCT Gazette.</i></p> |
|--|---|

(54) Title: HOMOGENISATE PHYTYL TRANSFERASE

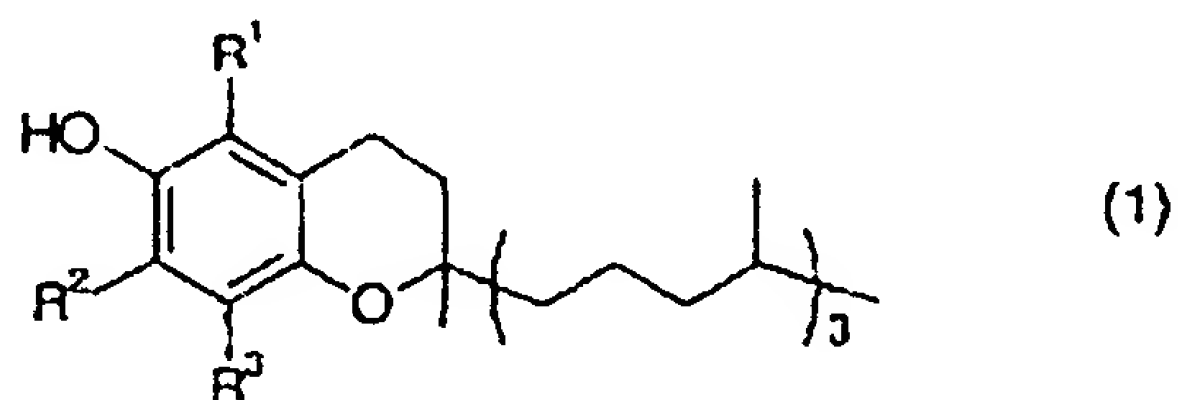
(57) Abstract: The invention relates to nucleic acid sequences that code for a protein with homogentisate phytyl transferase activity; to the use of these nucleic acid sequences for producing transgenic organisms such as transgenic plants with an increased tocopherol and tocotrienol content, to a method for producing plants with an increased tocopherol and/or tocotrienol content, and to the transgenic plants themselves.

## Homogentisate phytyl transferase

## Specification

The invention relates to nucleic acid sequences that code for a protein with homogentisate phytyl transferase activity, to the use of these nucleic acids for producing transgenic organisms such as for example transgenic plants with an increased content of tocopherols and/or tocotrienols, to a method for producing plants with an increased content of tocopherols and/or tocotrienols, and to the transgenic organisms, such as for example transgenic plants themselves.

The eight compounds with vitamin E activity that occur in nature are derivatives of 6-chromanol (Ullmann's Encyclopedia of Industrial Chemistry, Vol. A 27 (1996), VCH Verlagsgesellschaft, Chapter 4, 478-488, Vitamin E). The group of the tocopherols (1a-d) features a saturated side chain; the group of the tocotrienols (2a-d) features an unsaturated side chain:



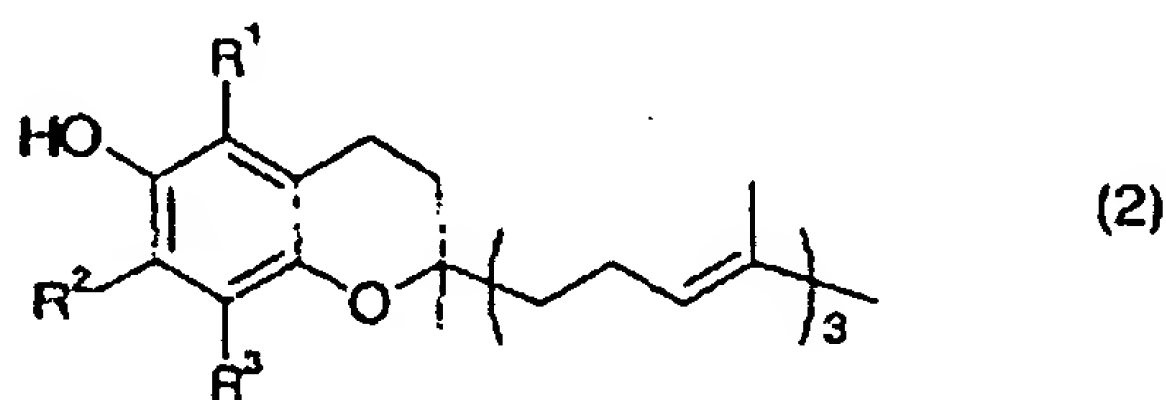
1a,  $\alpha$ -tocopherol:  $R^1 = R^2 = R^3 = \text{CH}_3$

1b,  $\beta$ -tocopherol [148-03-8]:  $R^1 = R^3 = \text{CH}_3$ ,  $R^2 = \text{H}$

1c,  $\gamma$ -tocopherol [54-28-4]:  $R^1 = \text{H}$ ,  $R^2 = R^3 = \text{CH}_3$

1d,  $\delta$ -tocopherol [119-13-1]:  $R^1 = R^2 = \text{H}$ ,  $R^3 = \text{CH}_3$

Insert formula 2



2a,  $\alpha$ -tocotrienol [1721-51-3]:  $R^1 = R^2 = R^3 = \text{CH}_3$

2b,  $\beta$ -tocotrienol [490-23-3]:  $R^1 = R^3 = \text{CH}_3$ ,  $R^2 = \text{H}$

2c,  $\gamma$ -tocotrienol [14101-61-2]:  $R^1 = \text{H}$ ,  $R^2 = R^3 = \text{CH}_3$

2d,  $\delta$ -tocotrienol [25612-59-3]:  $R^1 = R^2 = \text{H}$ ,  $R^3 = \text{CH}_3$

In the present invention, vitamin E is understood to mean all eight previously mentioned tocopherols and tocotrienols with vitamin E activity.

These compounds with vitamin E activity are important natural fat-soluble antioxidants. A lack of vitamin E leads to pathophysiological situations in humans and animals. Vitamin E compounds therefore have a high economic value as additives in the food and feed sector, in pharmaceutical formulations, and in cosmetic applications.

An economical method for producing vitamin E compounds, and food and feed with an increased vitamin E content, are thus extremely important.

Particularly economical methods are biotechnology methods that use proteins and biosynthesis genes of tocopherol or tocotrienol biosynthesis from vitamin E-producing organisms.

Figure 5 shows a biosynthesis diagram of tocopherols and tocotrienols.

In the course of the biosynthesis, homogentisic acid (homogentisate) is bound to phytyl pyrophosphate (PPP) or geranylgeranyl pyrophosphate in order to form the precursors of  $\alpha$ -tocopherol and  $\alpha$ -tocotrienol, 2-methylphytylhydroquinone or 2-methylgeranylgeranylhydroquinone respectively. First 2,3-dimethyl-6-phytylhydroquinone is formed by means of methylation steps with S-adenosylmethionine as a methyl group donor, then  $\gamma$ -tocopherol is formed by cyclization and  $\alpha$ -tocopherol by methylation once more.

Katani et al., Annu. Rev. Plant Physiol. Plant Mol. Biol., 1998, 49, 151 to 157, describe the total genomic sequence of the cyanobacterium *Synechocystis sp.* PCC6803.

Little is known as yet about increasing the flow of metabolites to increase the tocopherol or tocotrienol content in transgenic organisms, for example in transgenic plants, by overexpression of individual biosynthesis genes.

WO 97/27285 describes a modification of the tocopherol content by increased expression or by down-regulation of the enzyme p-hydroxyphenylpyruvate dioxygenase (HPPD).

WO 99/04622 describes gene sequences coding for a  $\gamma$ -tocopherol methyl transferase from *Synechocystis* PCC6803 and *Arabidopsis thaliana* and their incorporation into transgenic plants.

WO 99/23231 shows that the expression of a geranylgeranyl reductase in transgenic plants results in an increased tocopherol biosynthesis.

The object of the invention was to make available another biosynthesis gene of the vitamin E biosynthesis pathway and thus further advantageous transgenic plants with an increased content of tocopherols and tocotrienols.

The object was achieved by discovering nucleic acid sequences that code for a homogentisate phytyl transferase and by overexpression of the homogentisate phytyl transferase gene in plants.

Accordingly the present invention relates to proteins that feature the activity of a homogentisate phytyl transferase (HGPT), i.e. feature the ability to bind phytyl pyrophosphate to homogentisate, i.e. for example feature an enzymatic activity for the transformation of homogentisate and phytyl pyrophosphate into 2-methylphytylhydroquinone.

Preferred 2-methylphytylhydroquinones are 2-methyl-6-phytylhydroquinone or 2-methyl-5-phytylhydroquinone.

Homogentisate phytyl transferases are understood below to mean the proteins of the invention.

Preferred proteins feature the enzymatic activity for the transformation of homogentisate and phytyl pyrophosphate into 2-methylphytylhydroquinone and contain the amino acid sequence SEQ ID NO. 2 or a sequence derived from this sequence by substitution, insertion, or deletion of amino acids that features a homology of at least 20%, preferably 40%, preferably at least 60%, more preferred at least 80%, particularly preferred at least 90%, at the amino acid level with the sequence SEQ ID NO. 2.

Further examples of the proteins according to the invention can readily be discovered for example from various organisms whose genomic sequence is known, such as for example from *Arabidopsis thaliana* by homology comparisons of the amino acid sequences or the corresponding back-translated nucleic acid sequences from data banks with the SEQ ID NO. 2.

The proteins of the invention can be used as homogentisate phytyl transferases.

The preferred proteins are preferred for all uses according to the invention of the proteins according to the invention.

Substitution is understood to mean the replacement of one or more amino acids by one or more amino acids. It is preferred to carry out "conservative" replacements, in which the replacement amino acid has a property similar to that of the original amino acid, for example the replacement of Glu by Asp, Gln by Asn, Val by Ile, Leu by Ile, Ser by Thr.

Deletion is the replacement of an amino acid by a direct bond. Preferred positions for deletions are the termini of the polypeptide and the links between the individual protein domains.

Insertions are insertions of amino acids into the polypeptide chain, whereby theoretically a direct bond is replaced by one or more amino acids.

Homology between two proteins is understood to mean the identity of the amino acids over their respective total protein length, which is calculated by comparing them with the aid of the GAP program algorithm (UWGCG, University of Wisconsin, Genetic Computer Group) using the following parameters:

Gap Weight:	12
Length Weight:	4
Average Match:	2.912
Average Mismatch:	-2.003

Accordingly, a protein featuring a homology of at least 20% at the amino acid level with the sequence SEQ ID NO. 2 is understood to mean a protein that features a homology of at least 20% when its sequence is compared with the sequence SEQ ID NO. 2 according to the above program algorithm with the above set of parameters.

The homogentisate phytyl transferases according to the invention are capable of converting homogentisate derivatives and phytyl pyrophosphate derivatives into 2-methylphytylhydroquinone derivatives and/or homogentisate derivatives and geranylgeranyl pyrophosphate derivatives into 2-methylgeranylgeranylhydroquinone derivatives.

Homogentisate derivatives are understood to mean homogentisate and homogentisate compounds derived therefrom that are accepted as a substrate by the homogentisate phytyl transferases of the invention.

Phytyl pyrophosphate derivatives are understood to mean phytyl pyrophosphate and phytyl pyrophosphate compounds derived therefrom that are accepted as a substrate by the homogentisate phytyl transferases of the invention.

Accordingly, 2-methylphytylhydroquinone derivatives are understood to mean the compounds resulting from the enzymatic reaction, such as for example 2-methylphytylhydroquinone and the corresponding derived compounds.

Preferred 2-methylphytylhydroquinone derivatives are derivatives of 2-methyl-6-phytylhydroquinone or 2-methyl-5-phytylhydroquinone.

Geranylgeranyl pyrophosphate derivatives are understood to mean geranylgeranyl pyrophosphate and geranylgeranyl pyrophosphate compounds derived therefrom that are accepted as a substrate by the homogentisate phytyl transferases of the invention.

Accordingly, 2-methyl-geranylgeranylhydroquinone derivatives are understood to mean the compounds resulting from the enzymatic reaction, such as for example 2-methylgeranylgeranylhydroquinone and the corresponding derived compounds.

Preferred 2-methylgeranylgeranylhydroquinones are 2-methyl-6-geranylgeranylhydroquinone or 2-methyl-5-geranylgeranylhydroquinone.

Preferred 2-methylgeranylgeranylhydroquinone derivatives are derivatives of 2-methyl-6-geranylgeranylhydroquinone or 2-methyl-5-geranylgeranylhydroquinone.

Accordingly the invention relates to a method for the biotransformation, characterized in that homogentisate derivatives and phytyl pyrophosphate derivatives are converted into 2-methylphytylhydroquinone derivatives or homogentisate derivatives and geranylgeranyl pyrophosphate derivatives are converted into 2-methylgeranylgeranylhydroquinone derivatives in the presence of a homogentisate phytyl transferase of the invention.



In principle, the biotransformation can be carried out with whole cells that express the enzyme HGPT or cell extracts from these cells or else with purified or high-purity HGPT. The homogentisate phytyl transferase can also be in free or immobilized form thereby.

The homogentisate phytyl transferases according to the invention can also be used to produce vitamin E. The enzymatic biosynthesis step of the homogentisate phytyl transferases can take place thereby *in vitro* or as described below *in vivo*, for example in transgenic organisms, such as for example in transgenic plants.

Accordingly the invention relates to a method for the production of vitamin E, characterized in that homogentisate derivatives and phytyl pyrophosphate derivatives are converted into 2-methylphytylhydroquinone derivatives or homogentisate derivatives and geranylgeranyl pyrophosphate derivatives are converted into 2-methylgeranylgeranylhydroquinone derivatives in the presence of a homogentisate phytyl transferase of the invention.

The biosynthesis pathway of vitamin E moreover offers target enzymes for the development of inhibitors. Since according to the current state of the art no enzyme identical or similar to *Synechocystis* HGPT exists in human and animal organisms, it must be assumed that inhibitors act very specifically on plants.

Therefore the invention also relates to the use of the homogentisate phytyl transferase of the invention as a herbicidal target for the discovery of inhibitors of homogentisate phytyl transferase.

HGPT is a target for herbicides. In order to be able to find efficient inhibitors of HGPT, it is necessary to make available suitable test systems with which inhibitor-enzyme binding studies can be performed. For this purpose for example the complete cDNA sequence of HGPT from *Synechocystis* is cloned into an expression vector (pQE, Qiagen) and overexpressed in *E. coli*.

The HGPT protein expressed with the aid of the expression cassette of the invention is particularly suitable for discovering the HGPT-specific inhibitors.

Accordingly the invention relates to a method for discovering inhibitors of the homogentisate phytyl transferase, characterized in that the enzymatic activity of the homogentisate phytyl

transferase is measured in the presence of a chemical compound and in that if the enzymatic activity is reduced in comparison with the non-inhibited activity, the chemical compound represents an inhibitor.

In addition the HGPT can be used for example in an enzyme test in which the activity of the HGPT is determined in the presence and absence of the active substance to be tested. By comparing the two activity determinations, qualitative and quantitative information can be gained about the inhibitory behavior of the active substance to be tested.

A plurality of chemical compounds can be tested rapidly and simply for herbicidal properties with the aid of the test system of the invention. The method allows substances with a high level of activity to be selected reproducibly and specifically from a large number of substances, in order subsequently to carry out on these substances further in-depth tests familiar to those skilled in the art.

Another subject of the invention is therefore herbicidal active substances that can be identified with the above-described test system.

The homogenisate phytyl transferases according to the invention can be produced from natural or genetically altered organisms as described below, by means of gene expression of the corresponding nucleic acids that code for these proteins.

Another subject of the invention is nucleic acids, referred to below as homogenisate phytyl transferase genes (HPGT genes), that code for the previously described proteins of the invention.

The nucleic acid sequence can be for example an RNA, DNA, or cDNA sequence. For insertion into a nucleic acid construct, such as for example an expression cassette, suitable coding sequences are for example those that code for an HGPT and that endow the host with the capacity for overproduction of tocopherols and/or tocotrienols.

Suitable nucleic acid sequences are obtainable by backtranslation of the polypeptide sequence according to the genetic code.

For this, it is preferable to use codons that are frequently used in accordance with codon usage specific to the organism. The codon usage can readily be determined on the basis of computer evaluations of other known genes of the organism in question.



If the protein is to be expressed for example in a plant, it is frequently advantageous to use the codon usage of the plant in the backtranslation.

Preferred nucleic acids code for a plant homogentisate phytyl transferase or a homogentisate phytyl transferase from cyanobacteria.

A particularly preferred nucleic acid has the sequence SEQ ID NO. 1. This nucleic acid represents a prokaryotic genomic DNA from the cyanobacterium *Synechocystis* sp. PCC6803 that codes for the homogentisate phytyl transferase of the sequence SEQ ID NO. 2.

All the above-mentioned homogentisate phytyl transferase genes can be produced in a *per se* known manner by chemical synthesis from the nucleotide structural units such as for example by fragment condensation of individual overlapping complementary nucleic acid structural units of the double helix. The chemical synthesis of oligonucleotides can be carried out for example in a known manner by the phosphoramidite method (Voet, Voet, 2<sup>nd</sup> edition, Wiley Press New York, pp. 896-897). The attachment of synthetic oligonucleotides and filling of gaps with the aid of the Klenow fragment of the DNA polymerase and ligation reactions, as well as general cloning methods, are described in Sambrook et al. (1989), Molecular cloning: A laboratory manual, Cold Spring Harbor Laboratory Press.

The invention further relates to the use of the HGPT of the invention or the HGPT genes of the invention for the production of antibodies.

The invention further relates to nucleic acid constructs containing one of the above-described homogentisate phytyl transferase genes according to the invention that are functionally linked to one or more regulation signals that guarantee the transcription and translation in prokaryotic or eukaryotic organisms.

These regulatory sequences are for example sequences to which inducers or repressors bind and thus regulate the expression of the nucleic acid. In addition to these new regulation sequences or in place of these sequences, the natural regulation of these sequences upstream of the actual

structural genes can still be present and may possibly have been altered genetically, so that the natural regulation was switched off and the expression of the genes was increased. The nucleic acid construct can also be constructed more simply, however, i.e. no additional regulation signals are inserted before the above-mentioned homogentisate phytyl transferase genes and the natural promoter with its regulation is not removed. Instead, the natural regulation sequence is mutated so that regulation no longer occurs and the gene expression is increased. These changed promoters can also be brought alone upstream of the natural genes, to increase the activity.

Moreover the nucleic acid construct can advantageously also contain one or more "enhancer sequences" functionally linked with the promoter, that enable an increased expression of the nucleic acid sequence. Additional advantageous sequences, such as further regulatory elements or terminators, can also be inserted at the 3'-end of the DNA sequences. The above-mentioned homogentisate phytyl transferase genes can be contained in the gene construct in one or more copies.

It is preferred to use nucleic acid constructs that enable the expression of the homogentisate phytyl transferase gene of the invention in a host cell, also referred to below as an expression cassette.

The expression cassettes contain regulatory nucleic acid sequences that control the expression of the coding sequence in the host cell. According to a preferred form of embodiment, an expression cassette includes upstream, i.e. at the 5'-end of the coding sequence, a promoter, and downstream, i.e. at the 3'-end, a polyadenylation signal and optionally other regulatory elements functionally linked with the intermediate coding sequence for the homogentisate phytyl transferase gene.

A functional linking is understood to mean the sequential arrangement of promoter, coding sequence, terminator, and optionally other regulatory elements, such that each of the regulatory elements can fulfill its function in the expression of the coding sequence as intended. The sequences preferred but not limited thereto for the operative linking are targeting sequences to guarantee the subcellular localization in the apoplast, in the vacuoles, in plastids, in the

---

mitochondrion, in the endoplasmic reticulum (ER), in the cell nucleus, in the fat bodies or other compartments, and translation amplifiers such as the 5'-leading sequence from the tobacco mosaic virus (Gallie et al., Nucl. Acids Res. 15 (1987), 8693-8711).

Depending on the host organism or starting organism described in greater detail below that is converted into a genetically altered or transgenic organism by the incorporation of the expression cassette, various regulation sequences are suitable.

Advantageous regulation sequences for the nucleic acid constructs of the invention, for the method for the production of vitamin E described below, and for the genetically altered organisms described below, are contained for example in promoters such as cos-, tac-, trp-, tet-, trp-tet-, lpp-, lac-, lpp-lac-, lacIq-, T7-, T5-, T3-, gal-, trc-, ara-, SP6-, l-PR-, or the l-PL promoter, which are advantageously used in gram-negative bacteria.

Other advantageous regulation sequences are contained for example in the gram-positive promoters amy and SPO2, in the yeast or fungus promoters ADC1, MFa, AC, P-60, CYC1, GAPDH, TEF, rp28, ADH, or in the plant promoters CaMV/35S (Franck et al., Cell 21 (1980) 285-294), PRP1 (Ward et al., Plant. Mol. Biol. 22 (1993)), SSU, OCS, leb4, usp, STLS1, B33, nos, or in the ubiquitin or phaseolin promoter.

For plants as genetically altered organisms, in principle any promoter that can control the expression of foreign genes in plants, is suitable as promoter of the expression cassette.

It is preferable in particular to use a plant promoter or a promoter derived from a plant virus. The CaMV 35S promoter from the cauliflower mosaic virus (Franck et al., Cell 21 (1980), 285-294) is particularly preferred. As is known, this promoter contains various recognition sequences for transcriptional effectors that in their totality lead to a permanent and constitutive expression of the introduced gene (Benfey et al., EMBO J. 8 (1989), 2195-2202). The expression cassette can also contain a pathogen- or chemically inducible promoter through which the expression of the exogenous homogentisate phytyl transferase gene in the plant can be controlled at a certain point in time.

Such promoters as e.g. the PRP1 promoter (Ward et al., Plant. Mol. Biol. 22 (1993), 361-366), a promoter inducible by salicylic acid (WO 95/19443), a promoter inducible by benzene

---

sulfonamide (EP-A 388186), a promoter inducible by tetracycline (Gatz et al., (1992) Plant J. 2, 397-404), a promoter inducible by abscisic acid (EP-A 335528), or a promoter inducible by ethanol or cyclohexanone (WO 93/21334) can be used for example.

Moreover promoters are particularly preferred that ensure the expression in tissues or plant parts in which for example the biosynthesis of tocopherol or its precursors takes place or in which the products are advantageously accumulated.

Promoters to be named in particular are those for the whole plant based on constitutive expression, such as for example the CaMV promoter, the OCS promoter from *Agrobacterium* (octopine synthase), the NOS promoter from *Agrobacterium* (nopaline synthase), the ubiquitin promoter, promoters of vacuolar ATPase subunits or the promoter of a proline-rich protein from wheat (wheat WO 9113991).

Also to be named in particular are promoters that guarantee a leaf-specific expression. The promoter of the cytosolic FBPase from potato (WO 9705900), the SSU promoter (small subunit) of rubisco (ribulose 1,5-bisphosphate carboxylase) or the ST-LSI promoter from potato (Stockhaus et al., EMBO J. 8 (1989), 2445-245[sic]) are to be named.

Other suitable promoters are for example

specific promoters for tubers, storage roots, or roots, such as for example the patatin promoter class I (B33), the promoter of the cathepsin D inhibitor from potato, the promoter of starch synthesis (GBSS1) or the sporamin promoter,

fruit-specific promoters, such as for example the fruit-specific promoter from tomato (EP 409625),

fruit ripening-specific promoters, such as for example the fruit-ripening-specific promoter from tomato (WO 9421794),

flower-specific promoters, such as for example the phytoene synthase promoter (WO 9216635) or the promoter of the P-rr gene (WO 9822593) or

specific plastid- or chromoplast promoters such as for example the RNA polymerase promoter (WO 9706250) or

---

also the promoter of the phosphoribosyl pyrophosphate amidotransferase from *Glycine max* (see also Genbank Accession Number U87999) or another node-specific promoter as in EP 249676 can be used advantageously.

In principle, all natural promoters with their regulation sequences can be used like the above-named promoters for the method of the invention. Synthetic promoters can also be used advantageously.

As an example, the plant expression cassette can be inserted into a derivative of the transformation vector pBin19 with 35s promoter (Bevan, M., Nucleic Acids Research 12: 8711-8721 (1984)). Figure 2 shows a derivative of the transformation vector pBin-19 with seed-specific legumin B4 promoter.

The expression cassette can contain for example a seed-specific promoter (preferably the phaseolin promoter (US 5 504200), the USP promoter (Baumlein, H. et al., Mol. Gen. Genet. (1991) 225 (3), 459-467), the Bce4 gene promoter from *Brassica* WO 9113980) or LEB4 promoter (Fiedler and Conrad, 1995)), the LEB4 signal peptide, the gene to be expressed, and an ER retention signal.

An expression cassette is produced for example by fusion of a suitable promoter with a suitable HGPT DNA sequence and preferably a DNA inserted between promoter and HGPT DNA sequence, which DNA codes for a chloroplast-specific transit peptide, as well as with a polyadenylation signal according to current recombination- and cloning techniques, such as described for example in T. Maniatis, E. F. Fritsch and J. Sambrook, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1989) and in T.J. Silhavy, M.L. Berman and L.W. Enquist, Experiments with Gene Fusions, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1984) and in Ausubel, F.M. et al., Current Protocols in Molecular Biology, Greene Publishing Assoc. and Wiley-Interscience (1987).

Sequences that guarantee a targeting in the plastids are particularly preferred.

Expression cassettes can also be used whose DNA sequence codes for example for an HGPT fusion protein, whereby a part of the fusion protein is a transit peptide that controls the translocation of the polypeptide. Chloroplast-specific transit peptides are preferred that are

---



enzymatically cleaved from the HGPT part after translocation of the HGPT gene into the chloroplasts. The transit peptide derived from the plastidial *Nicotiana tabacum* transketolase or from another transit peptide (e.g. the transit peptide of the small subunit of rubisco or of the ferredoxin NADP oxidoreductase as well as of the isopentenyl pyrophosphate isomerase-2) or its functional equivalent is particularly preferred.

DNA sequences of three cassettes of the plastid-transit peptide of the plastidial transketolase of tobacco in three reading frames as KpnI/BamHI fragments with an ATG codon in the NcoI cleavage site are particularly preferred:

pTP09

KpnI\_GGTACCATGGCGTCTTCTTCTTCTCACTCTCTCTCAAGCTATCCTCTCTCGTT  
CTGTCCCTCGCCATGGCTCTGCCTCTTCTTCTCAACTTTCCCCTTCTTCTCACTTTT  
TCCGGCCTTAAATCCAATCCCAATATCACACCTCCCGCCGCGGTACTCCTTCCTCCG  
CCGCCGCGCGCGCGTCGTAAGGTCACCGGCGATTCTGTGCCTCAGCTGCAACCGAA  
ACCATAGAGAAAAGTGAAGTGCAGGATCC\_BamHI

pTP10

KpnI\_GGTACCATGGCGTCTTCTTCTTCTCACTCTCTCTCAAGCTATCCTCTCTCGTT  
CTGTCCCTCGCCATGGCTCTGCCTCTTCTTCTCAACTTTCCCCTTCTTCTCACTTTT  
TCCGGCCTTAAATCCAATCCCAATATCACACCTCCCGCCGCGGTACTCCTTCCTCCG  
CCGCCGCGCGCGCGTCGTAAGGTCACCGGCGATTCTGTGCCTCAGCTGCAACCGAA  
ACCATAGAGAAAAGTGAAGTGCAGGATCC\_BamHI

pTP11

KpnI\_GGTACCATGGCGTCTTCTTCTTCTCACTCTCTCTCAAGCTATCCTCTCTCGTT  
CTGTCCCTCGCCATGGCTCTGCCTCTTCTTCTCAACTTTCCCCTTCTTCTCACTTTT  
TCCGGCCTTAAATCCAATCCCAATATCACACCTCCCGCCGCGGTACTCCTTCCTCCG  
CCGCCGCGCGCGCGTCGTAAGGTCACCGGCGATTCTGTGCCTCAGCTGCAACCGAA  
ACCATAGAGAAAAGTGAAGTGCAGGATCC\_BamHI

The inserted nucleotide sequence coding for an HGPT can be produced synthetically or obtained naturally or can contain a mixture of synthetic and natural DNA constituents, as well as of different heterologous HGPT gene segments of different organisms. In general, synthetic nucleotide sequences are produced with codons that are preferred by plants. These codons preferred by plants can be determined from codons with the highest protein frequency that are expressed in the most important plant species. In the preparation of an expression cassette,



various DNA fragments can be manipulated in order to obtain a nucleotide sequence that expediently reads in the correct direction and that is equipped with a correct reading frame. Adapters or linkers can be attached to the fragments to join the DNA fragments together.

The promoter and terminator regions can expediently be provided in the transcription direction with a linker or polylinker that contains one or more restriction sites for the insertion of this sequence. As a rule the linker has 1 to 10, usually 1 to 8, preferably 2 to 6, restriction sites. In general the size of the linker within the regulatory regions is less than 100 bp, frequently less than 60 bp, but at least 5 bp. The promoter can be both native or homologous to the host plant, and foreign or heterologous to it. The expression cassette contains in the 5'-3'-transcription direction the promoter, a DNA sequence that codes for an HGPT gene, and a region for the transcriptional termination. Different termination regions can be substituted for one another as desired.

Manipulations that prepare suitable restriction cleavage sites or remove the superfluous DNA or restriction cleavage sites can also be used. As far as insertions, deletions, or substitutions such as e.g. transitions and transversions are concerned, *in vitro* mutagenesis, "primer repair," restriction, or ligation can be used. With suitable manipulations, such as e.g. restriction, "chewing back," or the filling up of overhangs for "blunt ends," complementary ends of the fragments can be made available for the ligation.

Preferred polyadenylation signals are plant polyadenylation signals, preferably those that essentially correspond to T-DNA polyadenylation signals from *Agrobacterium tumefaciens*, in particular of gene 3 of the T-DNA (octopine synthase) of the Ti plasmid pTiACH5 (Gielen et al., EMBO J. 3 (1984), 835 ff) or functional equivalents.

The fused expression cassette, which codes for an HGPT gene, is preferably cloned into a vector, for example pBin19, which is suitable for transforming *Agrobacterium tumefaciens*. *Agrobacteria* transformed with such a vector can then be used in a known manner for the transformation of plants, in particular crop plants, such as e.g. tobacco plants, for example by bathing injured leaves

---

or leaf pieces in an *Agrobacteria* solution and then cultivating them in suitable media. The transformation of plants by *Agrobacteria* is known i.a. from F.F. White, Vectors for Gene Transfer in Higher Plants; in Transgenic Plants, Vol. 1, Engineering and Utilization, edited by S.D. Kung and R. Wu, Academic Press, 1993, pp. 15-38. Transgenic plants that contain a gene for the expression of an HGPT gene integrated into the expression cassette, can be regenerated in a known manner from the transformed cells of the injured leaves or leaf pieces.

The nucleic acid constructs according to the invention can be used for the production of genetically altered organisms. The genetically altered organisms are produced by transformation of the host organisms, also referred to below as starting organisms, with a construct containing the HGPT gene.

Starting or host organisms are understood to mean prokaryotic or eukaryotic organisms such as for example microorganisms, mosses, or plants. Preferred microorganisms are bacteria, yeasts, algae, or fungi.

Preferred bacteria are bacteria of the genus *Escherichia*, *Erwinia*, *Agrobacterium*, *Flavobacterium*, *Alcaligenes*, or cyanobacteria of the genus *Synechocystis*.

Preferred yeasts are *Candida*, *Saccharomyces*, *Hansenula*, or *Pichia*.

Preferred fungi are *Aspergillus*, *Trichoderma*, *Ashbya*, *Neurospora*, *Fusarium*, or other fungi described in Indian Chem Engr. Section B. Vol 37, No 1,2 (1995) on page 15, Table 6.

Preferred algae are green algae, such as for example algae of the genus *Haematococcus*, *Phaedactylum tricornatum*, *Volvox*, or *Dunaliella*.

The invention relates to a genetically altered organism, whereby the genetic change

if the starting organism contains a nucleic acid of the invention, increases the gene expression of a nucleic acid of the invention compared with a wild type, or

if the starting organism does not contain a nucleic acid of the invention, causes the gene expression of a nucleic acid of the invention compared with a wild type.

---

The transgenic organisms containing the HGPT gene of the invention are capable of converting homogentisate derivatives and phytyl pyrophosphate derivatives into 2-methylphytylhydroquinone derivatives and/or homogentisate derivatives and geranylgeranyl pyrophosphate derivatives into 2-methylgeranylgeranylhydroquinone derivatives.

These organisms can be used for example for the above-described biotransformation.

Transgenic organisms containing an exogenous HGPT gene of the invention, which as starting organisms already possess the biosynthesis genes for the production of vitamin E, such as for example plants or other photosynthetically active organisms such as for example cyanobacteria, mosses, or algae, feature an increased content of tocopherols and/or tocotrienols in comparison with the respective wild type.

The invention therefore relates to such a genetically altered organism of the invention that features an increased vitamin E content compared with the wild type.

The present invention also relates to the use of the HGPT of the invention or the HGPT genes of the invention for the production of vitamin E in transgenic organisms.

Genetically altered organisms of the invention, preferably plants that feature an increased vitamin E content compared with the wild type, can be used for the production of vitamin E.

The present invention therefore also relates to methods for the production of vitamin E, in that a genetically altered organism of the invention, preferably a genetically altered plant of the invention, which features an increased vitamin E content compared with the wild type, is cultivated, the organism is harvested, and the vitamin E compounds are subsequently isolated from the organism.

Genetically altered plants of the invention with an increased vitamin E content that can be consumed by humans and animals can also be used as food or feed for example, directly or after a *per se* known preparation.

To produce organisms with an increased vitamin E content (tocopherols and/or tocotrienols) compared with the wild type, in a preferred form of embodiment plants are used as starting organisms and correspondingly also as genetically altered organisms.

---

Preferred plants are for example *Tagetes*, sunflower, *Arabidopsis*, tobacco, red pepper, soybean, tomato, eggplant, paprika, various types of carrot, potato, corn, lettuce and brassicaceous plants, cereals, alfalfa, oats, barley, rye, wheat, triticale, millet, rice, lucerne, flax, cotton, hemp, Brassicaceae such as for example rape or canola, sugar-beet, sugar-cane, nut and grape species, or woody plants such as for example aspen or yew.

*Arabidopsis thaliana*, *Tagetes erecta*, *Brassica napus*, *Nicotiana tabacum*, canola, potatoes and oilseeds, such as for example soybean, are particularly preferred.

The present invention furthermore relates to a method for the production of genetically altered organisms in which a nucleic acid of the invention or a nucleic acid construct of the invention is introduced into the genome of the starting organism.

For the transformation of a host organism, such as for example a plant, with a DNA coding for an HGPT, an expression cassette is incorporated as an insertion into a recombinant vector whose vector DNA can preferably contain additional functional regulation signals, for example sequences for replication or integration.

Suitable vectors for plants are described i.a. in "Methods in Plant Molecular Biology and Biotechnology" (CRC Press), Chapter 6-7, pp. 71-119 (1993).

Using the above-mentioned recombination and cloning techniques, the expression cassettes can be cloned into suitable vectors that enable them to multiply, for example into *E. coli*. Suitable cloning vectors are i.a. pBR332, pUC series, M13mp series and pACYC184. Binary vectors that can replicate in both *E. coli* and *Agrobacteria* are particularly suitable.

A further subject of the invention relates to the use of an expression cassette containing a DNA sequence SEQ ID NO. 1 or a DNA sequence hybridizing with this for the transformation of plants, plant cells, plant tissues, or plant parts. The goal of the use is preferably to increase the plant content of tocopherols and/or tocotrienols. Depending on the choice of promoter, the

---

expression can take place specifically in the leaves, the seeds, petals, or other parts of the plant. Such transgenic plants, their reproductive material, as well as their plant cells, plant tissues, or plant parts, are another subject of the present invention.

The expression cassette can moreover be used for the transformation of bacteria, cyanobacteria, yeasts, filamentous fungi, mosses, and algae, with the goal of increasing the content of tocopherols and/or tocotrienols.

The transfer of foreign genes into the genome of a plant is called transformation. The described methods for the transformation and regeneration of plants from plant tissues or plant cells are used thereby for the transient or stable transformation. Suitable methods are protoplast transformation by polyethylene glycol-induced DNA uptake, the biolistic method with the gene gun – the so-called particle bombardment method, electroporation, incubation of dry embryos in DNA-containing solution, microinjection, and gene transfer mediated by *Agrobacterium*. The said methods are described for example in B. Jenes et al., Techniques for Gene Transfer, in: Transgenic Plants, Vol. 1, Engineering and Utilization, edited by S.D. Kung and R. Wu, Academic Press (1993), 128-143, as well as in Potrykus, Annu. Rev. Plant Physiol. Plant Molec. Biol. 42 (1991), 205-225). Preferably the construct to be expressed is cloned in a vector that is suitable for transforming *Agrobacterium tumefaciens*, for example pBin19 (Bevan et al., Nucl. Acids Res. 12 (1984), 8711).

*Agrobacteria* transformed with an expression cassette can likewise be used in a known manner for the transformation of plants, e.g. by bathing injured leaves or leaf parts in an *Agrobacteria* solution and then cultivating them in suitable media.

Increasing the content of tocopherols or tocotrienols means, in the context of the present invention, the artificially acquired ability of an increased biosynthesis performance of these compounds by functional overexpression of an HGPT gene of the invention in the plant compared with the plant not modified by genetic engineering.

---

Both the content of tocopherols or tocotrienols can be increased thereby. Preferably the content of tocopherols is increased. However, it is also possible under certain conditions for the content of tocotrienols to be increased in preference.

The biosynthesis site of tocopherols for example is i.a. the leaf tissue, so that a leaf-specific expression of the HGPT gene is feasible. However, it is obvious that the tocopherol biosynthesis does not have to be limited to the leaf tissue, but can also take place tissue-specifically in all other parts of the plant – particularly in fat-containing seeds.

Moreover, a constitutive expression of the exogenous HGPT gene is advantageous. On the other hand, however, an inducible expression can also appear desirable.

The effectiveness of the expression of the transgenically expressed HGPT gene can be ascertained for example *in vitro* by shoot meristem proliferation. In addition, an expression of the HGPT gene changed in type and level and its effect on the tocopherol biosynthesis performance can be tested on test plants in greenhouse tests.

A subject of the invention is moreover transgenic plants transformed with an expression cassette containing an HGPT gene of the invention, as well as transgenic cells, tissue, parts, and reproductive material of such plants.

As mentioned above, transgenic plants such as for example *Tagetes*, sunflower, *Arabidopsis*, tobacco, red pepper, soybean, tomato, eggplant, paprika, various types of carrot, potato, corn, lettuce and brassicaceous plants, cereals, alfalfa, oats, barley, rye, wheat, triticale, millet, rice, lucerne, flax, cotton, hemp, Brassicaceae such as for example rape or canola, sugar-beet, sugar-cane, nut and grape species, or woody plants such as for example aspen or yew are preferred thereby.

Plants in the sense of the invention are monocotyledonous and dicotyledonous plants.

Another subject of the invention is further photosynthetically active organisms transformed with an expression cassette containing an HGPT gene of the invention.

In addition to the increased vitamin E content, an increased resistance to inhibitors of HGPT is also achieved in a plant by overexpression of the gene sequence coding for an HGPT of the invention.



Therefore the invention relates to a genetically altered organism of the invention, preferably a genetically altered plant of the invention, that features a resistance to inhibitors of the homogentisate phytyl transferase.

By means of the present invention it is possible for example in transgenic plants to successfully increase the activity of the homogentisate phytyl transferase (HGPT) by overexpression of the HGPT gene of the invention. This can be achieved in principle by expression of homologous or heterologous HGPT genes.

In Example 1, the cloning of an HGPT DNA sequence (SEQ ID NO. 1) from *Synechocystis* sp. PCC 6803 is described for the first time. In order to guarantee plastid localization, a transit signal sequence (Fig. 1-4) is placed upstream of the HGPT nucleotide sequence from *Synechocystis*.

The available 2-methylphytylhydroquinone or 2-methylgeranylgeranylhydroquinone now multiplied by means of the additional expression of the HGPT gene, is reacted further in the direction of tocopherols and tocotrienol (Figure 5).

Measurements on HGPT *Synechocystis* knock-out mutants showed a drastic decrease with respect to the content of tocopherols. This proves the direct influence of the plastidial plant HGPT on the synthesis of tocopherols and tocotrienols.

Further subjects of the invention are:

- Methods for the transformation of a plant, characterized in that expression cassettes containing an HGPT gene of the invention are incorporated into a plant cell or protoplasts of plants, and these are regenerated to produce whole plants.
- Use of the HGPT gene of the invention for the production of plants with an increased content of tocopherols and/or tocotrienols by expression of an HGPT DNA sequence in plants.

The invention is explained by the following Examples, but is not limited to these:

General conditions:

Sequence analysis of recombinant DNA

---

The sequencing of recombinant DNA molecules was performed with a Licor laser fluorescence DNA sequencer (sold through MWG Biotech, Ebersbach) by the method of Sanger (Sanger et al., Proc. Natl. Acad. Sci. USA 74 (1977), 5463-5467).

#### Example 1

Cloning of the homogentisate phytyl transferase from *Synechocystis sp.* PCC 6803.

The DNA coding for the ORF slr1736 was amplified by means of polymerase chain reaction (PCR) from *Synechocystis sp.* PCC 6803 according to the method of Crispin A. Howitt (BioTechniques 21:32-34, July 1996), using a sense specific primer (slr17365' Figure 8, SEQ ID NO. 3) and an antisense specific primer (slr17363', Figure 9, SEQ ID NO. 4).

The PCR conditions were as follows:

The PCR took place in a 50 µL reaction batch in which were contained:

- 5 µL of a *Synechocystis sp.* PCC 6803 cell suspension
- 0.2 mM of dATP, dTTP, dGTP, dCTP
- 1.5 mM of Mg (OAc)<sub>2</sub>
- 5 µg of bovine serum albumin
- 40 pmol of slr17365'
- 40 pmol of slr17363'
- 15 µL of 3.3 x rTth DNA polymerase XL buffer (PE Applied Biosystems)
- 5 U rTth DNA polymerase XL (PE Applied Biosystems)

The PCR was carried out under the following cycle conditions:

- Step 1: 5 minutes 94°C (denaturing)
- Step 2: 3 seconds 94°C
- Step 3: 2 minutes 48°C (annealing)
- Step 4: 2 minutes 72°C (elongation)
- 35 repetitions of Steps 2-4
- Step 5: 10 minutes 72°C (post-elongation)
- Step 6: 4°C (wait loop)

The amplicon was cloned in the PCR cloning vector pGEM-T (Promega), using standard methods. The identity of the amplicon produced was confirmed by sequencing using the M13F (-40) primer.

### Example 2

#### Production of an slr1736 knock-out mutant

A DNA construct for the production of a deletion mutant of the ORF slr1736 in *Synechocystis sp.* PCC 6803 was produced, using standard cloning techniques.

The vector pGEM-T/slr1736 was digested using restriction enzyme HpaI. An internal fragment of slr1736 including 348 bp was deleted by means of this digestion. Then the aminoglycoside-3'-phosphotransferase of the transposon Tn903 was cloned into the HpaI cleavage sites. For this, the Tn903 was isolated as an EcoRI fragment from the vector pUC4k (Vieira, J. and Messing, J., Gene: 19, 259-268, 1982), the projecting ends of the restriction digestion are converted into blunt ends by standard methods and ligated into the HpaI cut vector pGEM-T/slr1736. The ligation batch was used for the transformation of *E. coli* X11 blue cells. Transformants were selected by using kanamycin and ampicillin. A recombinant plasmid (pGEM-T/slr1736::tn903, see Fig. 6) was isolated and used for the transformation of *Synechocystis sp.* PCC 6803 according to the method of Williams (Methods Enzymol. 167:776-778, 1987).

Figure 6 shows a construct for the knock-out mutagenesis of the ORF slr1736 in *Synechocystis sp.* PCC 6803.

*Synechocystis sp.* PCC 6803 transformants were selected on kanamycin-containing (km) BG-11 solid medium (Castenholz, Methods in Enzymology, 68-93, 1988) at 28°C and 30  $\mu\text{mol photons} \times (\text{m}^2 \times \text{s})^{-1}$ . Four independent knock-out mutants were able to be produced after five selection rounds (transfers of individual colonies to fresh BG-11 km medium).

The complete loss of the slr1736 endogene or its exchange for the recombinant slr1736::tn903 DNA was confirmed by PCR analyses.

### Example 3

#### Comparison of tocopherol production in *Synechocystis sp.* PCC 6803 wild type cells and the produced knock-out mutants of the ORF slr1736

The cells of the four independent *Synechocystis sp.* PCC 6803 knock-out mutants of the ORF slr1736 cultivated on the BG-11 km agar medium, as well as untransformed wild type cells, were used to inoculate liquid cultures. These cultures were cultivated at 28°C and 30  $\mu\text{mol photons} \times$

$(\text{m}^2 \times \text{s})^{-1}$  ( $30 \mu\text{E}$ ) for about 3 days. After the  $\text{OD}_{730}$  of the individual cultures had been determined, the  $\text{OD}_{730}$  of all cultures was synchronized by means of corresponding dilutions with BG-11 (wild types) or BG-11 km (mutants). These cultures synchronized for cell density were used to inoculate three cultures per mutant or the wild type controls. The biochemical analyses were thus performed using respectively three independently grown cultures of a mutant and the corresponding wild types. The cultures were grown to an optical density of  $\text{OD}_{730} = 0.3$ . The medium of the cell culture was removed by centrifugation at 14 000 rpm carried out twice in an Eppendorf table centrifuge. The subsequent lysis of the cells was accomplished by means of incubation four times in the Eppendorf shaker at  $30^\circ\text{C}$ , 1000 rpm in 100% methanol for 15 minutes, whereby the respectively obtained supernatants were combined. Further incubation steps gave no further liberation of tocopherols or tocotrienols.

In order to avoid oxidation, the obtained extracts were analyzed directly after the extraction with the aid of a Waters Alliance 2690 HPLC system. Tocopherols and tocotrienols were separated via a reverse phase column (ProntoSil 200-3-C30, Bischoff) with a mobile phase of 100% methanol and were identified based on standards (Merck). The fluorescence of the substances (excitation 295 nm, emission 320 nm), which was detected with the aid of a Jasco fluorescence detector FP 920, served as the detection system.

No tocopherols could be found in the *Synechocystis sp.* PCC 6803 knock-out mutants of the ORF slr1736. Tocopherols were measured in the *Synechocystis sp.* PCC 6803 wild type cells, however.

The loss of the ability to produce tocopherols within the knock-out mutants of the ORF slr1736 in comparison with the *Synechocystis sp.* PCC 6803 wild type cells shows that the slr1736 gene codes for a homogentisate phytyl transferase.

#### Example 4

Functional characterization of the homogentisate phytyl transferase from *Synechocystis sp.* PCC 6803 by heterologous expression in *E. coli*

The hypothetical protein slr1736 from *Synechocystis sp.* PCC 6803 was able to be identified as homogentisate phytyl transferase by functional expression in *E. coli*.

---

The gene slr1736 amplified from *Synechocystis* sp. PCC 6803 was subcloned in the correct reading frame in the expression vector pQE-30 (Qiagen). The primers slr17365' or slr17363' (SEQ ID NO. 2 and 3) used for the amplification of the ORF slr1736 from *Synechocystis* sp. PCC 6803 were constructed so that BamHI restriction cleavage sites were added at the 5' end and the 3' end of the amplicon, see SEQ ID NO. 1. The slr1736 fragment was isolated using these flanking BamHI restriction cleavage sites from the recombinant plasmid pGEM-T/slr1736 and was ligated into a BamHI cleaved pQE-30 using standard methods. The ligation batch was used for the transformation of M15 *E. coli* cells and kanamycin and ampicillin-resistant transformants were analyzed. The kanamycin resistance is transmitted by the pREP-4 plasmid contained in the M15 cells. A recombinant plasmid (pQE-30/slr1736) that carried the slr1736 fragment in the correct orientation was isolated. The identity and orientation of the insert was confirmed by sequencing.

The recombinant plasmid pQE-30/slr1736 was used for the transformation of M15 *E. coli* cells in order to produce recombinant slr1736 protein. Using a colony originating from the transformation, an overnight culture in Luria broth medium was inoculated with 200 µg/mL of ampicillin (amp) and 50 µg/mL of kanamycin (km). Starting from this culture, a 100 mL Luria broth culture (amp/km) was inoculated the following morning. This culture was incubated at 28°C on a shaking incubator until an OD<sub>600</sub>: 0.35-0.4 was reached. Then the production of the recombinant protein was induced by adding 0.4 mM of isopropyl-β-D-thiogalactopyranoside (IPTG). The culture was shaken at 28°C for a further 3 hours and the cells were then pelleted by centrifugation at 8000 g.

The pellet was resuspended in 600 µL of lysis buffer (about 1-1.5 mL/g pellet wet weight, 10 mM HEPES KOH pH 7.8, 5 mM of dithiothreitol (DTT), 0.24 M sorbitol). Then PMSF (phenylmethyl sulfonate) was added to a final concentration of 0.15 mM and the batch was placed on ice for 10 minutes. The cells were lysed by a 10-second ultrasound pulse using an ultrasound wand. After the addition of Triton X100 (final concentration 0.1%), the cell suspension was incubated on ice for 30 minutes. The batch was then centrifuged off at 25 000 xg for 30 minutes and the supernatant was used for the assay.



The activity of the homogentisate phytyl transferase was determined by detection of radioactively labeled 2-methylphytylhydroquinone as a reaction product.

Then 235  $\mu\text{L}$  of the enzyme (about 300-600  $\mu\text{g}$ ) together with 35  $\mu\text{L}$  of phytyl pyrophosphate and 50  $\mu\text{L}$  (1.2 nmol) of  $^3\text{H}$ -homogentisic acid was incubated at 25°C for 4 hours in the following reaction buffer: 100  $\mu\text{L}$  (250 mM) of tricine-NaOH pH 7.6, 100  $\mu\text{L}$  (1.25 mM) of sorbitol, 10  $\mu\text{L}$  (50 mM) of  $\text{MgCl}_2$ , and 20  $\mu\text{L}$  (250 mM) of ascorbate. The tritium-labeled homogentisic acid [was] present in an ethanolic solution with 1 mg of ascorbate/mL. 50  $\mu\text{L}$  of this was concentrated by evaporation and the buffer was added, as well as the enzyme and the phytyl pyrophosphate.

The reaction was stopped by extraction of the batch twice with ethyl acetate. The ethyl acetate phases were concentrated by evaporation and the residues were picked up in methanol and applied to a thin-layer plate for chromatographic separation of the substances (solid phase: HPTLC plates: silica gel 60 F<sub>254</sub> (Mer[c]k), liquid phase: toluene). The radioactively labeled reaction product is detected using a phospho imager.

These experiments confirmed that the protein coded for by the gene slr1736 (SEQ ID NO. 1) from *Synechocystis sp.* PCC 6803 is a homogentisate phytyl transferase, since it has the enzymatic activity to form 2-methylphytylhydroquinone from homogentisates and phytyl pyrophosphate.

#### Example 5

##### Production of expression cassettes containing the HGPT gene

Transgenic plants were produced that express the homogentisate phytyl transferase from *Synechocystis sp.* PCC 6803 on the one hand under the control of the constitutive 35S promoter of the CaMV (cauliflower mosaic virus) (Franck et al., Cell 21: 285-294, 1980) and on the other hand under the control of the seed-specific promoter of the legumin gene from *Vicia faba* (Kafatos et al., Nuc. Acid. Res., 14(6): 2707-2720, 1986). The basis of the plasmid produced from *Synechocystis sp.* PCC 6803 for the constitutive expression of the homogentisate phytyl transferase was the pBinAR-TkTp-9 (Ralf Badur, Dissertation Universität Göttingen, 1998). This vector is a derivative of the pBinAR (Höfgen and Willmitzer, Plant Sci. 66: 221-230, 1990) and contains the 35S promoter of the CaMV (cauliflower mosaic virus) (Franck et al., 1980), the termination signal of the octopine synthase gene (Gielen et al., EMBO J. 3: 835-846, 1984), as well as the DNA sequence coding for the transit peptide of the *Nicotiana tabacum* plastidial



transketolase. Cloning of the homogentisate phytyl transferase from *Synechocystis sp.* PCC 6803 into this vector, which took place taking into consideration the correct reading frame, produces a translation fusion of the homogentisate phytyl transferase with the plastidial transit peptide. This results in a transport of the transgene into the plastids.

To produce this plasmid, the gene slr1736 was isolated from the plasmid pGEM-T/slr1736 using the flanking BamHI restriction cleavage sites. Using standard methods, this fragment was ligated into a BamHI cleaved pBinAR-TkTp-9 (see Figure 1). This plasmid (pBinAR-TkTp-9/slr1736) was used for the production of transgenic *Nicotiana tabacum* plants.

Fragment A (529 bp) in Figure 1 contains the 35S promoter of the CaMV (nucleotides 6909 to 7437 of the cauliflower mosaic virus), fragment B (245 bp) fragment[sic] codes for the transit peptide of the *Nicotiana tabacum* transketolase, fragment C (944 bp) codes for ORF slr1736 from *Synechocystis sp.* PCC 6803, fragment D (219 bp) codes for the termination signal of the octopine synthase gene.

To produce a plasmid that enables the seed-specific expression of the homogentisate phytyl transferase from *Synechocystis sp.* PCC 6803 in plants, the seed-specific promoter of the legumin B4 gene (Kafatos et al., Nuc. Acid. Res., 14(6): 2707-2720, 1986) was used. The 2.7 Kb fragment of the legumin B4 gene promoter was isolated from the pCR script/lePOCS plasmid, using the EcoRI cleavage site flanking the promoter 5' and the KpnI cleavage site flanking the 3'. The plasmid pBinAR-TkTp-9/slr1736 was likewise treated with the restriction enzymes EcoRI and KpnI. The result of this was that the 35S promoter of the CaMV was separated out from this plasmid. The promoter of the legumin gene was then cloned into this vector as an EcoRI/KpnI fragment, so that a plasmid was produced that placed the expression of the gene slr1736 under the control of this seed-specific promoter, see Figure 2. This plasmid (pBinARleP-TkTp-9/slr1736) was used to produce transgenic *Nicotiana tabacum* plants.

Fragment A (2700 bp) in Figure 2 contains the promoter of the legumin B4 gene from *Vicia faba*, fragment B (245 bp) codes for the transit peptide of the *Nicotiana tabacum* transketolase, fragment C (944 bp) codes for the ORF slr1736 from *Synechocystis sp.* PCC 6803, fragment D (219 bp) codes for the termination signal of the octopine synthase gene.

Production of DNA constructs for the expression of the homogentisate phytyl transferase from *Synechocystis sp.* PCC 6803 in *A. thaliana* and *B. napus*.

To produce chimeric DNA constructs for the production of transgenic *A. thaliana* or *B. napus* plants that express the homogentisate phytyl transferase from *Synechocystis sp.* PCC 6803, the vectors pPTVkan35S-IPP-Tp-9OCS or pPTVkanLeP-IPP-Tp-10NOS were used. These vectors are derivatives of the pGPTVkan (D. Becker, E. Kemper, J. Schell, R. Masterson. *Plant Molecular Biology* 20: 1195-1197, 1992) from which the *uidA* gene was deleted. In its place, the pPTVkan35S-IPP-Tp-9OCS contains the 35S promoter of the CaMV (cauliflower mosaic virus) (Franck et al., 1980), the sequence coding for the transit peptide of the *A. thaliana* plastid-specific isopentenyl pyrophosphate isomerase-2 (IPP-2) (Badur, unpublished), and the termination signal of the octopine synthase gene (Gielen et al., 1984).

The vector pPTVkanLeP-IPP-Tp-10nos contains the seed-specific promoter of the legumin B4 gene (Kafatos et al., Nuc. Acid. Res., 14(6): 2707-2720, 1986), likewise the sequence coding for the transit peptide of the *A. thaliana* plastid-specific isopentenyl pyrophosphate isomerase-2 (IPP-2) (Badur, unpublished), and the termination signal of the nopaline synthase from *A. tumefaciens* (Depicker et al., *J. Mol. Appl. Genet.* 1, 561-73, 1982).

The DNA molecules coding for the ORF slr1736 from *Synechocystis sp.* PCC 6803 were cloned as BamHI fragments with blunt ends filled by the T4 polymerase into the vectors pPTVkan35S-IPP-Tp-9OCS (Fig. 3) or pPTVkanLeP-IPP-Tp-10nos (Fig. 4), so that a translation fusion with the transit peptide of the IPP-2 was produced. Thus an import of the homogentisate phytyl transferase into the plastids could be guaranteed.

Fragment A (529 bp) in Figure 3 contains the 35S promoter of the CaMV (nucleotides 6909 to 7437 of the cauliflower mosaic virus). Fragment B (205 bp) fragment coding for the transit peptide of the *A. thaliana* isopentenyl pyrophosphate isomerase-2. Fragment C (944 bp) ORF slr1736 from *Synechocystis sp.* PCC 6803. Fragment D (219 bp) termination signal of the octopine synthase gene.

Fragment A (2700 bp) in Figure 4 contains the promoter of the legumin B4 gene from *Vicia faba*, fragment B (206 bp) fragment coding for the transit peptide of the *A. thaliana* isopentenyl pyrophosphate isomerase-2, fragment C (944 bp) codes for the ORF slr1736 from *Synechocystis sp.* PCC 6803. Fragment D (272 bp) for the termination signal of the nopaline synthase gene.

#### Example 6

##### Production of transgenic *Arabidopsis thaliana* plants

Wild type *Arabidopsis thaliana* plants (Columbia) were transformed with the *Agrobacterium tumefaciens* strain (GV3101 (pMP901)) based on a modified vacuum infiltration method (Steve Clough and Andrew Bent. Floral dip: a simplified method for *Agrobacterium* mediated transformation of *A. thaliana*. Plant J 16 (6): 735-43, 1998; der Bechtold, N. Ellis, J. and Pelltier, G., in: Planta *Agrobacterium*-mediated gene transfer by infiltration of adult *Arabidopsis thaliana* plants. CRAcad Sci Paris, 1993, 1144(2): 204-212). The *Agrobacterium tumefaciens* cells used were transformed in advance with the plasmids pPTVkan35S-IPP-Tp-9/slr1736 or pPTVkanLeP-IPP-Tp-9/slr1736 (Figures 3 and 4).

Seeds of the primary transformants were selected on the basis of their antibiotic resistance. Antibiotic resistant seedlings were planted in soil, and as fully developed plants were used for biochemical analysis.

#### Example 7

##### Production of transgenic *Brassica napus* plants

The production of transgenic rape plants followed a protocol of Bade, J.B. and Damm, B. (in Gene Transfer to Plants, Potrykus, I. and Spangenberg, G., eds, Springer Lab Manual, Springer Verlag, 1995, 30-38), in which the composition of the media and buffers used is also given.

The transformations took place with the *Agrobacterium tumefaciens* strain GV3101 (pMP90). The plasmids pPTVkan35S-IPP-Tp-9/slr1736 or pPTVkanLeP-IPP-Tp-10/slr1736 (Figures 3 and 4) were used for the transformation. The surface of seeds of *Brassica napus* var. Westar was sterilized with 70% ethanol (v/v), the seeds were washed in water at 55°C for 10 minutes, incubated in 1% hypochlorite solution (25% v/v Teepol, 0.1% v/v Tween 20) for 20 minutes, and washed six times with sterile water for 20 minutes respectively. The seeds were dried on filter paper for three days and 10-15 seeds were placed in a glass flask with 15 mL germination medium for germination. The roots and apices were removed from several seedlings (about 10 cm in height), and the remaining hypocotyls were cut into pieces about 6 mm long. The approx. 600 explantates obtained in this manner were washed with 50 mL of basal medium for 30 minutes and

transferred into a 300-mL flask. After 100 mL of callus induction medium had been added, the cultures were incubated for 24 hours at 100 rpm.

An overnight culture of the *Agrobacterium* strain was made up in Luria broth medium with kanamycin (20 mg/L) at 29°C, 2 mL of which was incubated in 50 mL of Luria broth medium without kanamycin at 29°C for 4 hours, up to an OD<sub>600</sub> of 0.4-0.5. After pelleting of the culture at 2000 rpm for 25 min, the cell pellet was resuspended in 25 mL of basal medium. The concentration of the bacteria in the solution was adjusted to an OD<sub>600</sub> of 0.3 by adding further basal medium.

The callus induction medium was removed from the rape explants with sterile pipettes, 50 mL of *Agrobacterium* solution was added, mixed carefully, and incubated for 20 min. The *Agrobacteria* suspension was removed, the rape explants were washed with 50 mL of callus induction medium for 1 min, and then 100 mL of callus induction medium was added. The co-cultivation was carried out on a rotary shaker at 100 rpm for 24 h. The co-cultivation was stopped by removing the callus induction medium and the explants were washed twice at 100 rpm with 25 mL of washing medium for 1 min respectively and twice with 100 mL of washing medium respectively for 60 min. The washing medium with the explants was transferred into 15 cm Petri dishes and the medium was removed with sterile pipettes.

For the regeneration, 20-30 explants respectively were transferred into 90 mm Petri dishes that contained 25 mL of shoot induction medium with kanamycin. The Petri dishes were sealed with 2 layers of Leukopor and incubated at 25°C and 2000 lux for photoperiods of 16 hours light/8 hours darkness. Every 12 days the developing calli were moved to fresh Petri dishes with shoot induction medium. All the other steps for the regeneration of whole plants were carried out as described by Bade, J.B. and Damm, B. (in: Gene Transfer to Plants, Potrykus, I. and Spangenberg, G., eds, Springer Lab Manual, Springer Verlag, 1995, 30-38).

#### Example 8

Production of transgenic *Nicotiana tabacum* plants.

Ten mL of YEB medium with antibiotics (5 g/L of bovine extract, 1 g/L of yeast extract, 5 g/L of peptone, 5 g/L of saccharose, and 2 mM of MgSO<sub>4</sub>) was inoculated with a colony of *Agrobacterium tumefaciens* and cultivated overnight at 28°C. The cells were pelleted in a table centrifuge, 3500 rpm at 4°C for 20 min and were then resuspended in fresh YEB medium without antibiotics under sterile conditions. The cell suspension was used for the transformation.

The wild type plants from sterile culture were obtained by vegetative replication. For this, only the tip of the plant was cut off and transferred into a sterile preserving jar onto fresh 2MS medium. The hairs on the upper side of the leaves and the midribs of the leaves were removed from the remainder of the plant. The leaves were cut with a razor blade into pieces about 1 cm<sup>2</sup>. The *Agrobacteria* culture was transferred to a small Petri dish (diameter 2 cm). The leaf pieces were drawn through this solution briefly and laid with the leaf underside on 2MS medium in petri dishes (diameter 9 cm) so that they touched the medium. After two days in the dark at 25°C, the explants were transferred onto plates with callus induction medium and acclimated in the climatic chamber to 28°C. The medium had to be changed every 7-10 days. As soon as calli formed, the explants were transferred into sterile preserving jars on shoot induction medium with Claforan (0.6% BiTec agar (w/v), 2.0 mg/L of zeatin ribose, 0.02 mg/L of naphthylacetic acid, 0.02 mg/L of gibberelic acid, 0.25 g/mL of Claforan, 1.6% of glucose (w/v), and 50 mg/L of kanamycin). Organogenesis commenced after about one month and the shoots formed could be cut off. The shoots were cultivated on 2MS medium with Claforan and selection label. As soon as a vigorous root ball had formed, the plants could be potted in Pikier soil.

#### Example 9

##### Characterization of the transgenic plants

In order to confirm that the vitamin E biosynthesis in the transgenic plants is influenced by the expression of the homogentisate phytyl transferase from *Synechocystis* sp. PCC 6803, the tocopherol and tocotrienol contents in the leaves and seeds of the plants transformed with the described constructs (*Arabidopsis thaliana*, *Brassica napus*, and *Nicotiana tabacum*) are analyzed. For this, the transgenic plants are cultivated in the greenhouse and plants that express the gene coding for the homogentisate phytyl transferase from *Synechocystis* sp. PCC 6803 are analyzed at the Northern level. The tocopherol content and the tocotrienol content in the leaves and seeds of these plants is determined. In all cases the tocopherol or tocotrienol concentration in transgenic plants that additionally express a nucleic acid of the invention is higher than in non-transformed plants.



## Patent Claims

1. Protein that features an enzymatic activity for the transformation of homogentisate and phytyl pyrophosphate into 2-methylphytylhydroquinone.
2. Protein according to Claim 1, containing the amino acid sequence SEQ ID NO. 2 or a sequence derived from this sequence by substitution, insertion, or deletion of amino acids, that features a homology of at least 20% at the amino acid level with the sequence SEQ ID NO. 2.
3. Nucleic acid coding for a protein according to one of Claims 1 or 2.
4. Nucleic acid according to Claim 3, coding for a protein from plants, cyanobacteria, mosses, or algae.
5. Nucleic acid according to Claim 3, characterized in that it comprises the sequence represented in SEQ ID NO. 1.
6. Nucleic acid construct containing a nucleic acid according to one of Claims 3 through 5, that are functionally linked with one or more regulation signals that guarantee the transcription and translation in prokaryotic or eukaryotic organisms.
7. Nucleic acid construct according to Claim 6, characterized in that the regulation signals contain one or more promoters that guarantee the transcription and translation in prokaryotic or eukaryotic organisms.
8. Genetically altered organisms, whereby the genetic alteration
  - if the starting organism contains a nucleic acid according to one of Claims 3 to 5, increases the gene expression of a nucleic acid according to one of Claims 3 to 5, compared with a wild type, or
  - if the starting organism does not contain a nucleic acid according to one of Claims 3 to 5, causes the gene expression of a nucleic acid according to one of Claims 3 to 5, compared with a wild type.
9. Genetically altered organism according to Claim 8, characterized in that the genetically altered organism features an increased vitamin E content compared with the wild type.
10. Genetically altered organism according to Claim 8, characterized in that the genetically altered organism features a resistance to inhibitors of the homogentisate phytyl transferase.
11. Genetically altered organism according to one of Claims 8 to 10, characterized in that a eukaryotic organism is used as the organism.



12. Genetically altered organism according to Claim 11, characterized in that a plant is used as the eukaryotic organism.
13. Use of a genetically altered organism according to one of Claims 8 to 12 for the production of vitamin E or for the biotransformation of homogentisate derivatives and phytyl pyrophosphate derivatives into 2-methylphytylhydroquinone derivatives or homogentisate derivatives and geranylgeranyl pyrophosphate derivatives into 2-methylgeranylgeranylhydroquinone derivatives.
14. Method for the production of genetically altered organisms according to one of Claims 8 to 12, characterized in that a nucleic acid according to one of Claims 3 to 5 or a nucleic acid construct according to Claim 6 or 7 is introduced into the genome of the starting organism.
15. Use of the nucleic acid according to one of Claims 3 to 5 or of the nucleic acid constructs according to Claim 6 or 7 for the production of genetically altered organisms.
16. Use of the nucleic acid according to one of Claims 3 to 5 for the expression in organisms.
17. Method for the production of vitamin E, characterized in that homogentisate derivatives and phytyl pyrophosphate derivatives are converted into 2-methylphytylhydroquinone derivatives or homogentisate derivatives and geranylgeranyl pyrophosphate derivatives are converted into 2-methylgeranylgeranylhydroquinone derivatives in the presence of a protein according to one of Claims 1 or 2.
18. Method according to Claim 17, characterized in that an organism according to one of Claims 8 to 12 is cultivated, the organism is harvested, and the vitamin E compounds are subsequently isolated from the organism.
19. Method for the biotransformation, characterized in that homogentisate derivatives and phytyl pyrophosphate derivatives are converted into 2-methylphytylhydroquinone derivatives or homogentisate derivatives and geranylgeranyl pyrophosphate derivatives are converted into 2-methylgeranylgeranylhydroquinone derivatives in the presence of a protein according to one of Claims 1 or 2.
20. Method according to Claim 19, characterized in that the protein is present in an organism in free or immobilized form.
21. Use of the protein according to Claim 1 or 2 as a homogentisate phytyl transferase.
22. Use of the protein according to Claim 1 or 2 for the transformation of homogentisate derivatives and phytyl pyrophosphate derivatives into 2-methylphytylhydroquinone derivatives or homogentisate derivatives and geranylgeranyl pyrophosphate derivatives into 2-methylgeranylgeranylhydroquinone derivatives.

23. Use of the protein according to Claim 1 or 2 or of the nucleic acid according to one of Claims 3 to 5 for the production of vitamin E.
24. Use of the protein according to Claim 1 or 2 or of the nucleic acid according to one of Claims 3 to 5 for the production of vitamin E in transgenic organisms.
25. Use of the protein according to Claim 1 or 2 or of the nucleic acid according to one of Claims 3 to 5 for the production of vitamin E in transgenic plants.
26. Use of the protein according to Claim 1 or 2 or of the nucleic acid according to one of Claims 3 to 5 for the production of antibodies.
27. Use of the protein according to Claim 1 or 2 or of the nucleic acid according to one of Claims 3 to 5 as a herbicidal target for discovering inhibitors of the homogentisate phytyl transferase.
28. Method for discovering inhibitors of the homogentisate phytyl transferase, characterized in that the enzymatic activity of the homogentisate phytyl transferase is measured in the presence of a chemical compound and in that if the enzymatic activity is reduced in comparison with the non-inhibited activity, the chemical compound represents an inhibitor.
29. Herbicidal active substances, obtainable by a method according to Claim 28.

Figure 1

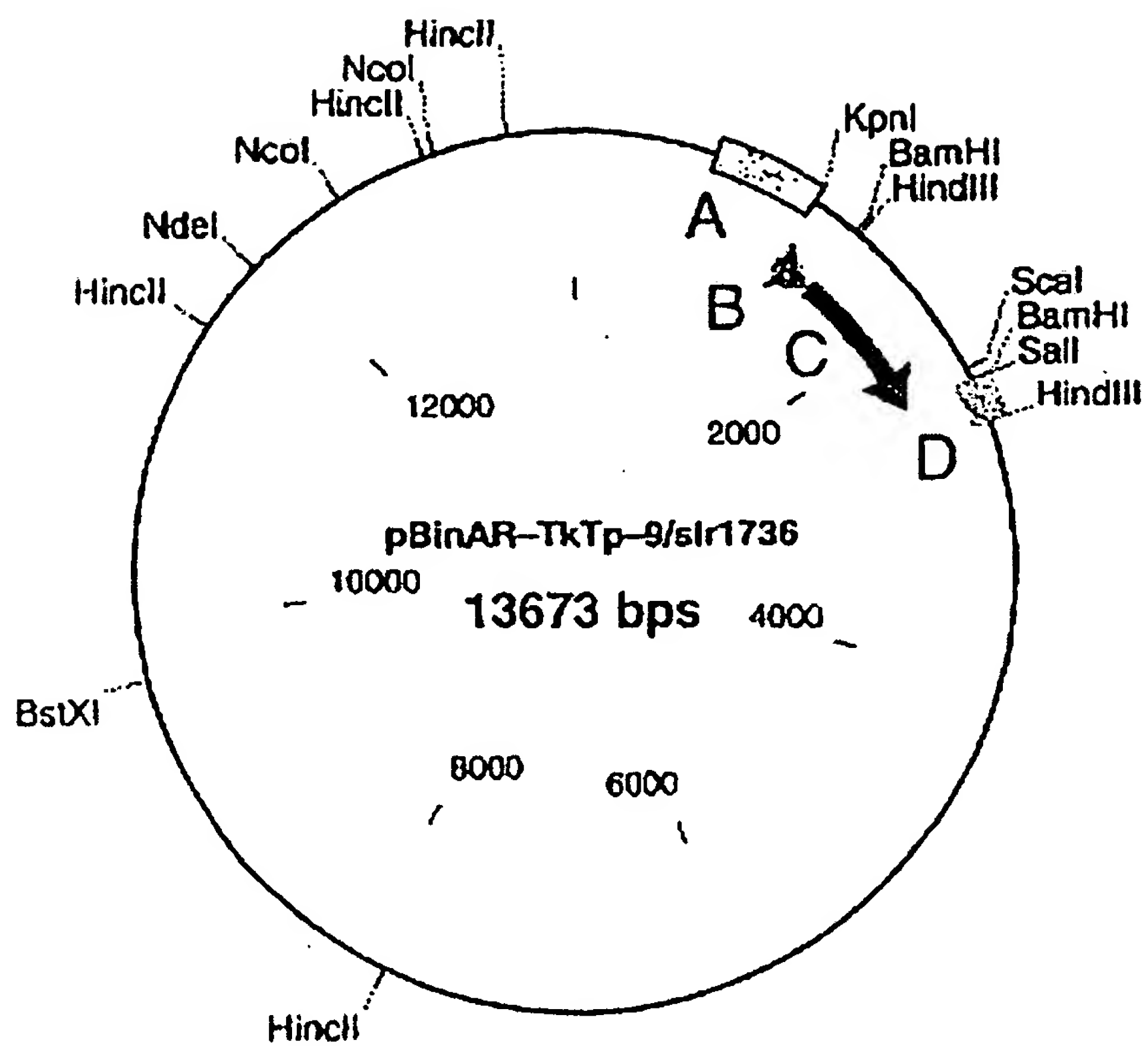


Figure 2

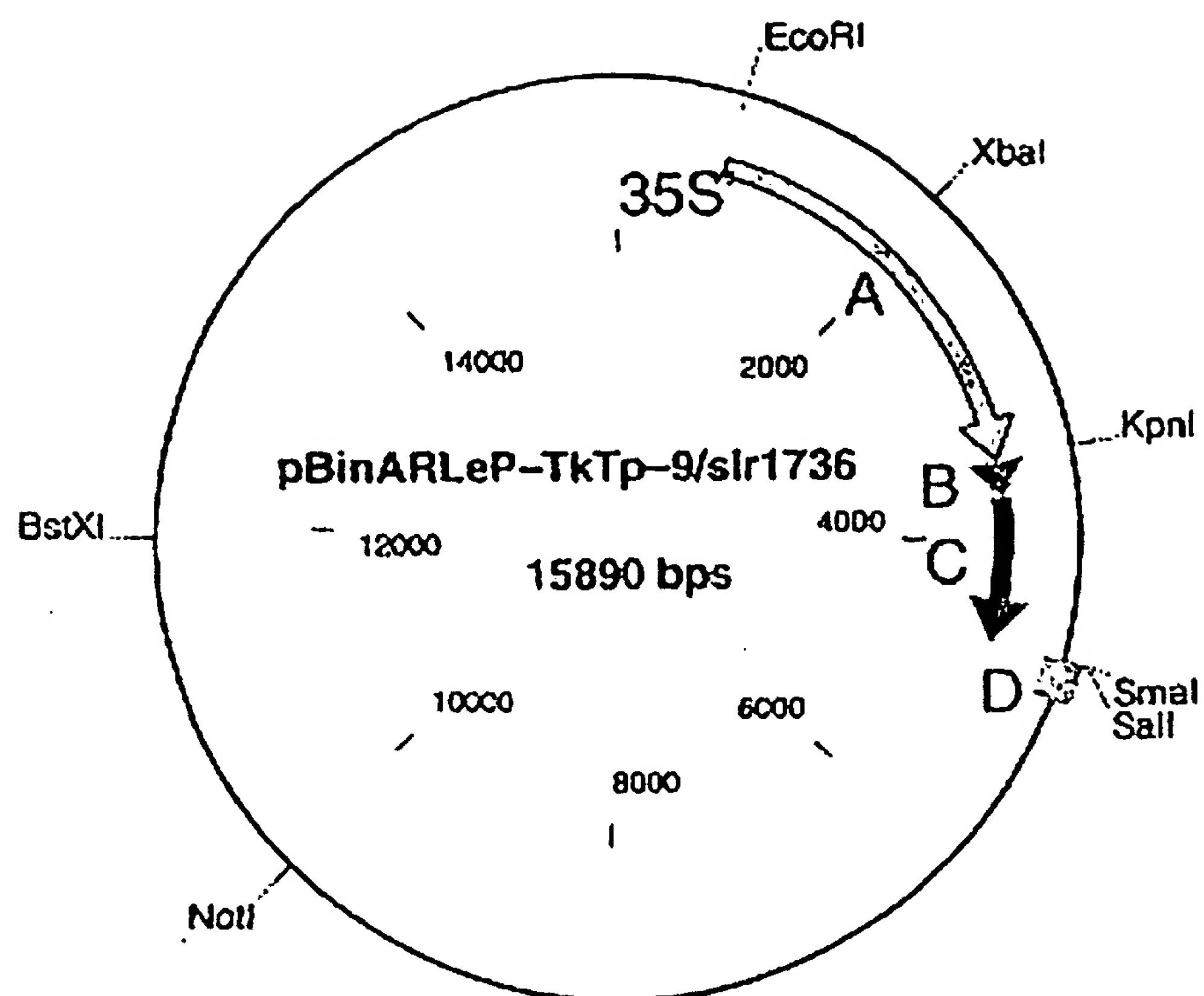


Figure 3

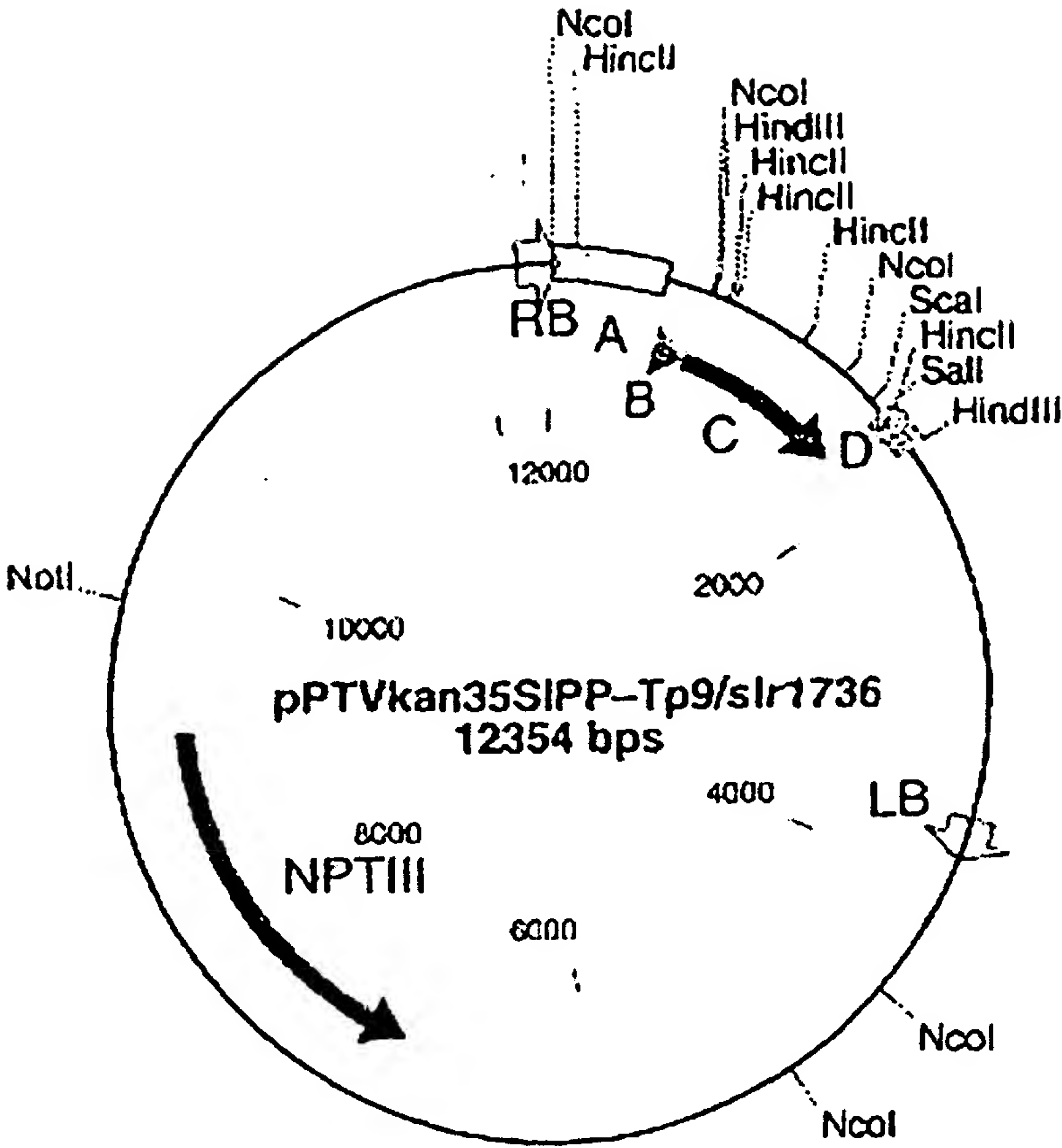


Figure 4

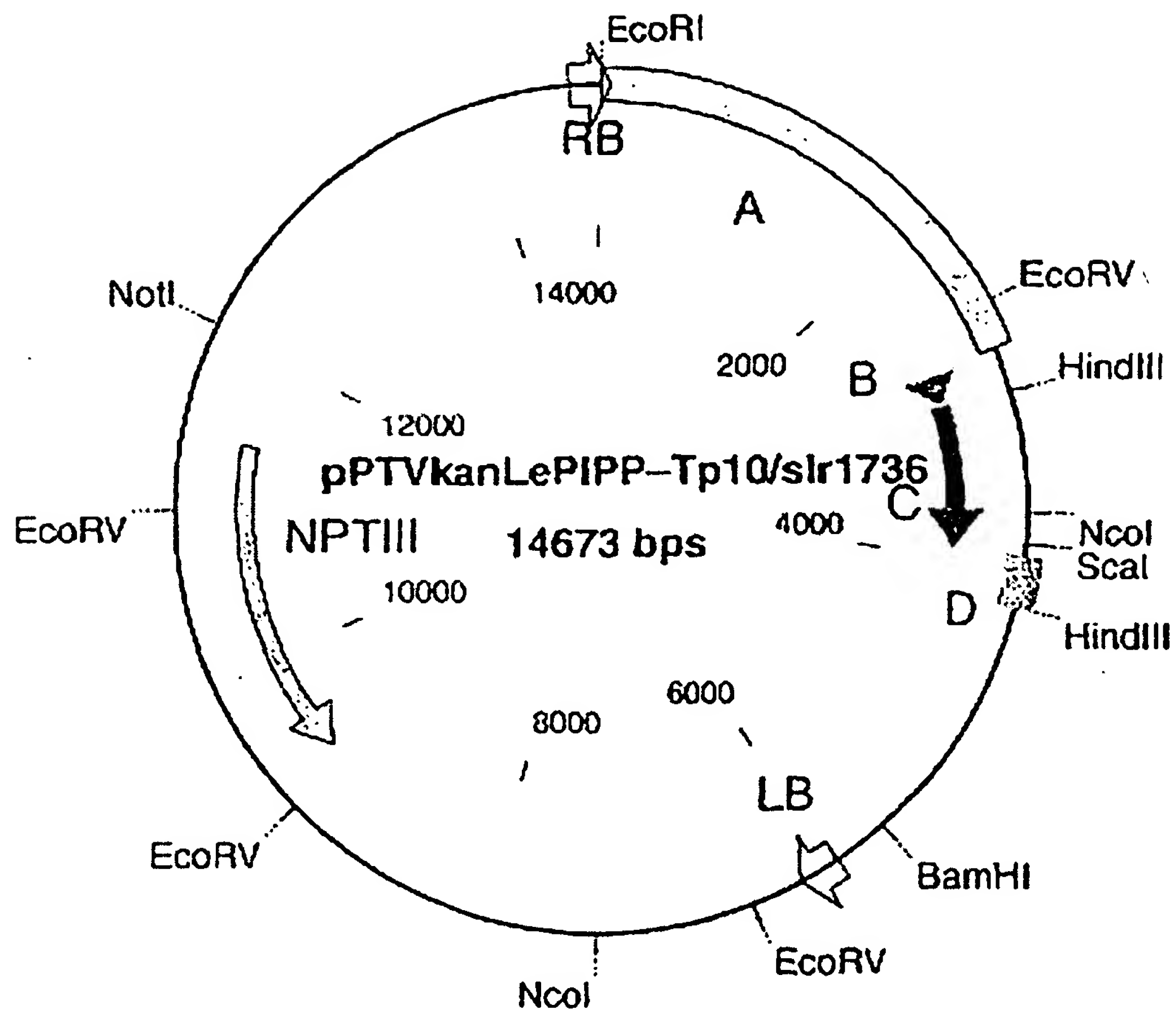




Figure 5

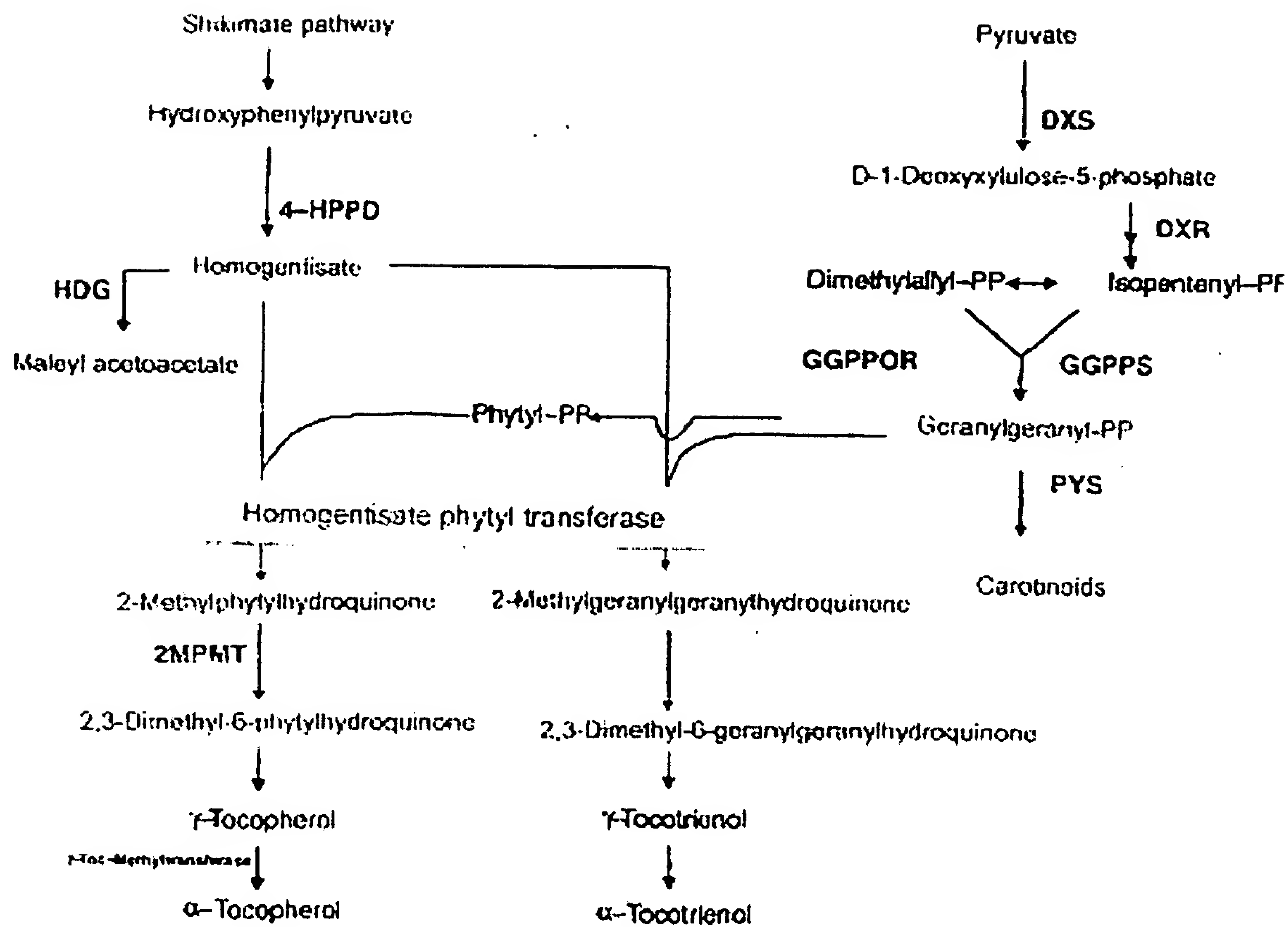


Figure 6

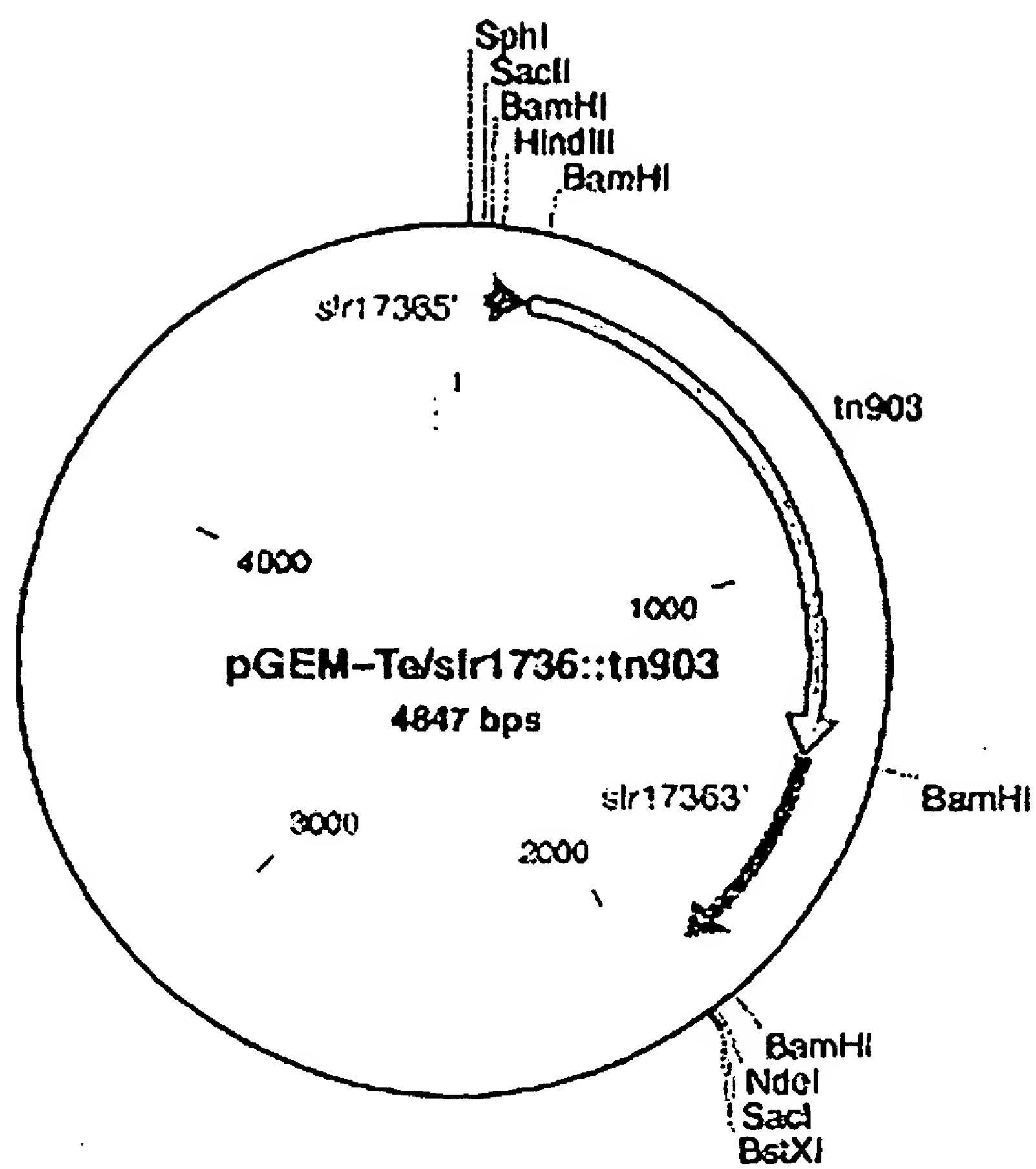


Figure 7

Sequence description: slr1736

Sequence characteristics:

Length: 944 bp

Type: Nucleic acid

Topology: Linear

Description: ORF of the hypothetical protein slr1736 from *Synechocystis* sp. PCC 6803

Upper-case letters: Sequence of the ORF slr1736

Lower-case letters: BamHI restriction cleavage sites added by PCR

```
ggaatccgccaatgccaaatatacgaacttttttggcgcttctcccccataccatcattgggtacaaactctgagcctctt  
ggctgtgpatctgtttaactattctggggatggaaactcagtttaactcccttgccttccctggatttatgtgttccgcgctt  
ccctgccttgcctgttcttaattgtttacattttccgcccacaaccaattttgggaatctogacattgaccccatcgaataag  
ccgaatttcccccctagctaacgagatgtttttctatcccccacggccctttggaattctggaacttttctggaattgttccctt  
gggaatcccttgggaatttgggctatgtctccgctaacccgtgggcaattacttttctttatctgacacgacctattccgtgc  
ccgcaatgagctttaaagcttctttccctgccttgcctcccttctttatcttccacggctccgggaattcttctttaaacttgc  
ttatttttatttttttgaatttgggtttagggttatcccccacttttaataaoccccatctcccttttgcactttatttatctt  
agttttcacccgtggcgatcgocatttttaagatgttgocagataiggaagocgatcgccattttaataatcacaacttttaa  
ctttgcaaatccgcaaaacaaacgtttttccgggaacttaattttactcactggttgttatttagccatggcgaatcttg  
cgcttatggcgacctatctctttaaatactcctttcttgatgttttccattttgtgcttatatgaccttactctggtggcg  
gaatcgacatgtacacttaaaagcacaacccgaattgctagtgtttatcagcttatttgcgaacctattttctttagagt  
acttgcgtatcccttgcctctgtggttacctaaatttttctaatactatttttttagggggaatcc
```

## Figure 8

Sequence description: slr17365'

Sequence characteristics:

Length: 26 bp

Type: Nucleic acid

Topology: Linear

Description: Oligonucleotide

5' -GGATCCGCCATGGCAACTATCCAAGC -3'

## Figure 9

Sequence description: slr17363'

Sequence characteristics:

Length: 25 bp

Type: Nucleic acid

Topology: Linear

Description: Oligonucleotide

5' -GGATCCCCCTAAAAAATAGTATTAG -3'

## SEQUENCE PROTOCOL

&lt;110&gt; SunGene GmbH &amp; Co KGAA

&lt;120&gt; Homogentisate phenyl transferase

&lt;130&gt; 0817/00012

&lt;140&gt;

&lt;141&gt;

&lt;160&gt; 4

&lt;170&gt; PatentIn Vers. 2.0

&lt;210&gt; 1

&lt;211&gt; 932

&lt;212&gt; DNA

&lt;213&gt; Synechocystis PCC6803

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (4)..(930)

&lt;400&gt; 1

```

gcc atg gca act atc caa gct ttt tgg cgc ttc tcc cgc ccc cat acc 48
Met Ala Thr Ile Gln Ala Phe Trp Arg Phe Ser Arg Pro His Thr
      1              5              10              15

atc att ggt acc act ctg agc gtc tgg gct gtg tat ctg tta act att 96
Ile Ile Gly Thr Thr Leu Ser Val Trp Ala Val Tyr Leu Leu Thr Ile
              20              25              30

ccc ggg gat gga aac tca gtt aac tcc cct gct tcc ctg gat tta gtg 144
Leu Gly Asp Gly Asn Ser Val Asn Ser Pro Ala Ser Leu Asp Leu Val
              35              40              45

ttc ggc gct tgg ctg gcc tgc ctg ttg ggt aac gtg tac att gtc ggc 192
Phe Gly Ala Trp Leu Ala Cys Leu Leu Gly Asn Val Tyr Ile Val Gly
              50              55              60

ctc aac caa ttg tgg gat gtg gac att gac cgc atc aat aag ccg aat 240
Leu Asn Gln Leu Trp Asp Val Asp Ile Asp Arg Ile Asn Lys Pro Asn
              65              70              75

ttg ccc cta gct aac gga gat ttt tct atc gcc cag ggc cgt tgg att 288
Leu Pro Leu Ala Asn Gly Asp Phe Ser Ile Ala Gln Gly Arg Trp Ile
              80              85              90              95

gtg gga cat tgt agc gtt gct tcc ttg gcg atc gcc tgg gga tta ggg 336

```



Val Gly Leu Cys Gly Val Ala Ser Leu Ala Ile Ala Trp Gly Leu Gly	
100 105 110	
cta tgg ctg ggg cta acg gtg ggc att agt ttg att att ggc acg gcc	384
Leu Trp Leu Gly Leu Thr Val Gly Ile Ser Leu Ile Ile Gly Thr Ala	
115 120 125	
tat tgg gtg ccg cca gtg agg tta aag cgc ttt tcc ctg ctg gcg gcc	432
Tyr Ser Val Pro Pro Val Arg Leu Lys Arg Phe Ser Leu Leu Ala Ala	
130 135 140	
ctg tgt att ctg acg gtg cgg gga att gtg gtt aac ttg ggc cta ttt	480
Leu Cys Ile Leu Thr Val Arg Gly Ile Val Val Asn Leu Gly Leu Phe	
145 150 155	
cta ttt ttt aga att ggt tta ggt tat ccc ccc act tta ata acc ccc	528
Leu Phe Phe Arg Ile Gly Leu Gly Tyr Pro Pro Thr Leu Ile Thr Pro	
160 165 170 175	
aac tgg gtt ttg act tta ttt atc tta gtt ttc acc gtg gcg atc gcc	576
Ile Trp Val Leu Thr Leu Phe Ile Leu Val Phe Thr Val Ala Ile Ala	
180 185 190	
att ttt aaa gat gtg cca gat atg gaa ggc gat cgg caa ttt aag att	624
Ile Phe Lys Asp Val Pro Asp Met Glu Gly Asp Arg Gln Phe Lys Ile	
195 200 205	
caa act tta acc ttg caa atc ggc aaa caa aac gtt ttt cgg gga acc	672
Gln Thr Leu Thr Leu Gln Ile Gly Lys Gln Asn Val Phe Arg Gly Thr	
210 215 220	
tta att tta ctc act ggt tgt tat tta gcc atg gca atc tgg ggc tta	720
Leu Ile Leu Leu Thr Gly Cys Tyr Leu Ala Met Ala Ile Trp Gly Leu	
225 230 235	
tgg gcy gct atg cct tta aat act gct ttc ttg att gtt tcc cat ttg	768
Trp Ala Ala Met Pro Leu Asn Thr Ala Phe Leu Ile Val Ser His Leu	
240 245 250 255	
tgc tta tta gcc tta ctc tgg tgg cgg agt cga gat gta cac tta gan	816
Cys Leu Leu Ala Leu Leu Trp Trp Arg Ser Arg Asp Val His Leu Glu	
260 265 270	
agc aaa acc gaa aat gct agt ttt tat cag ttt att tgg aag cta ttt	864
Ser Lys Thr Glu Ile Ala Ser Phe Tyr Gln Phe Ile Trp Lys Leu Phe	
275 280 285	
ttc tta gag tac ttg ctg tat ccc ttg gct ctg tgg tta cct aat ttt	912
Phe Leu Glu Tyr Leu Leu Tyr Pro Leu Ala Leu Trp Leu Pro Asn Phe	
290 295 300	

tct aat act att ttt tag gg  
 Ser Asn Thr Ile Phe  
 305

932

<210> 2  
 <211> 308  
 <212> PRT  
 <213> Synechocystis PCC6803

<400> 2  
 Met Ala Thr Ile Gln Ala Phe Trp Arg Phe Ser Arg Pro His Thr Ile  
 1 5 10 15  
 Ile Gly Thr Thr Leu Ser Val Trp Ala Val Tyr Leu Leu Thr Ile Leu  
 20 25 30  
 Gly Asp Gly Asn Ser Val Asn Ser Pro Ala Ser Leu Asp Leu Val Phe  
 35 40 45  
 Gly Ala Trp Leu Ala Cys Leu Leu Gly Asn Val Tyr Ile Val Gly Leu  
 50 55 60  
 Asn Gln Leu Trp Asp Val Asp Ile Asp Arg Ile Asn Lys Pro Asn Leu  
 65 70 75 80  
 Pro Leu Ala Asn Gly Asp Phe Ser Ile Ala Gln Gly Arg Trp Ile Val  
 85 90 95  
 Gly Leu Cys Gly Val Ala Ser Leu Ala Ile Ala Trp Gly Leu Gly Leu  
 100 105 110  
 Trp Leu Gly Leu Thr Val Gly Ile Ser Leu Ile Ile Gly Thr Ala Tyr  
 115 120 125  
 Ser Val Pro Pro Val Arg Leu Lys Arg Phe Ser Leu Leu Ala Ala Leu  
 130 135 140  
 Cys Ile Leu Thr Val Arg Gly Ile Val Val Asn Leu Gly Leu Phe Leu  
 145 150 155 160  
 Phe Phe Arg Ile Gly Leu Gly Tyr Pro Pro Thr Leu Ile Thr Pro Ile  
 165 170 175  
 Trp Val Leu Thr Leu Phe Ile Leu Val Phe Thr Val Ala Ile Ala Ile  
 180 185 190  
 Phe Lys Asp Val Pro Asp Met Glu Gly Asp Arg Gln Phe Lys Ile Gln  
 195 200 205

Thr Leu Thr Leu Gln Ile Gly Lys Gln Asn Val Phe Arg Gly Thr Leu  
210 215 220

Ile Leu Leu Thr Gly Cys Tyr Leu Ala Met Ala Ile Trp Gly Leu Trp  
225 230 235 240

Ala Ala Met Pro Leu Asn Thr Ala Phe Leu Ile Val Ser His Leu Cys  
245 250 255

Leu Leu Ala Leu Leu Trp Trp Arg Ser Arg Asp Val His Leu Glu Ser  
260 265 270

Lys Thr Glu Ile Ala Ser Phe Tyr Gln Phe Ile Trp Lys Leu Phe Phe  
275 280 285

Leu Glu Tyr Leu Leu Tyr Pro Leu Ala Leu Trp Leu Pro Asn Phe Ser  
290 295 300

Asn Thr Ile Phe  
305

<210> 3

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<221> primer\_bind

<222> (1)..(26)

<220>

<223> Description of Artificial Sequence: Primer

<400> 3

ggatccgcc tggcaactat ccagc

26

<210> 4

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<221> primer\_bind

<222> (1)..(25)

<220>

<223> Description of Artificial Sequence: Primer

<400> 4

ggatcccccct aaagaaatagt attag

25

Translation: GRIFFITHS, Technical Translator for GABT  
27 November 2002